Response Surface Methodology: An Analytical Method For Locating Migrated Contaminant Sources

THESIS

Orlando J. Dona Jr., 1st Lt, USAF

AFIT/ENG/GCS/97M

DEPARTMENT OF THE AIR FORCE
AIR UNIVERSITY
AIR FORCE INSTITUTE OF TECHNOLOGY

Wright-Patterson Air Force Base, Ohio

AFIT/GCS/ENC/97M-01

Response Surface Methodology: An Analytical Method For Locating Migrated Contaminant Sources

THESIS

Orlando J. Dona Jr., 1st Lt, USAF

AFIT/ENG/GCS/97M

THESIS APPROVAL

Student: 1st Lt Orlando J. Dona, Jr. Class: GCS-97M

Thesis Title: Response Surface Methodology: An Analytical Method For Locating Migrated Contaminant Sources

Defense Date: 21 Jan 97

Committee	Name/Department	Signature
Advisor:	Lt Col David Coulliette, Ph.D. Department of Mathematics and Statistics	
Co-Advisor:	Maj Edward Heyse, Ph.D. Department of Engineering and Environmental Management	
Reader:	Martin DeSimio, Ph.D. Department of Electrical and Computer Engineering	

DISCLAIMER

The views expressed in this thesis are those of the author
and do not reflect the official policy or position
of the Department of Defense or U. S. Government.

AFIT/GCS/ENG/97M-01

RESPONSE SURFACE METHODOLOGY: AN ANALYTICAL METHOD FOR LOCATING MIGRATED CONTAMINANT SOURCES

THESIS

Presented to the Faculty of the Graduate School of Engineering

of the Air Force Institute of Technology

Air Education and Training Command

In Partial Fulfillment of the

Requirements for the Degree of

Master of Science in Computer Systems

Orlando J. Dona Jr., B.A., M.S.

First Lieutenant, USAF

January, 1997

Acknowledgments

Many thanks to my advising team, Lt Col David Coulliette and Major Ed Heyse, for their more than generous support in this effort to integrate statistical methods and environmental engineering, both of which were fairly foreign to my background. Through their excellent direction, I have been able to contribute something of great worth to the Air Force and global community. Thanks to my neighbor and colleague Alex Hinojos, who helped me generate statistical data with short notice. Thanks to Marty who proof-read this document with such short notice. Special thanks to my wonderful spouse Patricia, darlings Denise and Ana (all 10 lb. 6 oz. birth weight of you, born just before the *Defense* of this thesis!) I owe you all my greatest appreciation and admiration for your patient enduring of the lost weekends and many after-work hours of study. Without your help and support, I could not have accomplished this great task. Thank you for keeping me focused on the important things in life. Finally, to my Lord and Savior Jesus Christ, without whose great gifts of Salvation, family, country and profession I would not be who and what I am today; thank You for Your help when I really needed it. I pray that You never let me forget that *worldly* wisdom will become as foolishness before You. Let this thesis and all the work that went into it be accomplished for Your glory -- in this world and the next.

OJ

Table of Contents

List of Figures

Figure Page

Figure Page

List of Tables

Abstract

The use and storage of hazardous chemicals at U.S. military facilities often adversely affect the groundwater when contaminants infiltrate the subsurface as a result of leaks and accidental spills. These contaminants, if not located and remediated in a relatively short time, may move and settle unpredictably, essentially creating a source some distance from the original leak or spill. An example of this phenomenon is found with migrating dense nonaqueous phase liquid (DNAPL) contaminants. Although various methods for estimating the present-day locations of these migrated contaminants are in use, accurately pinpointing the source of contaminants remains a difficult problem in current remediation technology. Response Surface Methodology (RSM) is a computer-enhanced statistical technique for empirical model building and exploitation that supports a systematic approach to site characterization. The use of RSM techniques may result in better mathematical models of a site and may ultimately enhance a site's conceptual model. This work demonstrates the use of RSM to pinpoint the statistically best locations of contaminant point sources that have migrated from their original location in several experiments, and outlines a process that has great potential for significantly reducing costs associated with site characterization and remediation.

I. Background

Many hazardous chemicals are stored and used at U.S. military facilities world-wide. These chemicals, including aircraft fuels, chemicals used in aircraft and motor vehicle maintenance and repair, and industrial wastewater, can infiltrate the subsurface as a result of leaks and accidental spills (Unger, 1995).

The cleanup, or remediation of these facilities, where necessary, is very costly to the DoD. In the 1996 budget, approximately $2.1 billion was earmarked for cleanup of DoD sites alone (money in addition to Superfund funding). This funding, which allows for the continued identification, investigation, and cleanup of past contamination leaks and spills, is an increase in funding of $500 million from 1995 (U.S. Budget, 1995). With increased congressional emphasis on across-the-board reduction of spending however, and especially in light of the large U.S. deficit, the DoD and other agencies can expect to have funding levels for environmental programs reduced considerably.

Traditional approaches to site remediation also produce a large amount of uncertainty. Where many potential sources of contamination exist in a large area, the ability to ascertain the actual source through sampling is an impossible task since the sample volume obtained from measurement wells is far too small. Increasing the amount of sampling is also unfeasible since the cost for additional samples adds very little information about the contaminated site. This uncertainty contributes to continuing

unsatisfactory long term results from remediation treatments based upon this limited sampling. There are also a number of research efforts which are currently attempting to tackle the migrated contaminant problem. However, these efforts are theoretical and have yet to mature for use outside an academic setting.

To overcome potential funding reductions and to make positive progress in site cleanup, new and alternative methods for site remediation must be researched and developed. This research work will demonstrate the application of Response Surface Methodology (RSM, a statistical technique for obtaining the best parameter inputs given known outputs) toward the site remediation problem, and will outline the steps required to employ RSM as a mathematical tool in determining the location and mass flux value of a contaminant source.

II. Literature Review and Research Goals

Introduction

A discussion of groundwater and source contaminants, site characterization and the conceptual model, mathematical modeling, and the inverse problem provide a basis for understanding RSM and its application to the groundwater and contaminant transport problem.

Groundwater and Source Contaminants

Many chemicals used on Air Force bases can pose a danger to human health even when the amount of subsurface infiltration is very small. Contaminants may migrate unpredictably, potentially impacting groundwater some distance from the original leak or spill. An example of this phenomenon is migrating dense nonaqueous phase liquid (DNAPL) contaminants, which because of their toxicity, limited solubility, and great migration potential in groundwater, contribute to serious long-term contamination (Unger, 1995; Cohen, 1993).

Although inexact methods are currently in use for localizing migrated contaminant sources, such as the EPA-recommended *phased approach* to contaminant plume mapping (EPA, 1989), obtaining a clear picture of the contaminated site remains a difficult problem in current remediation technology. The placement of pump-and-treat wells, for example, requires exact knowledge of the location of contaminant sources for remediation to be

successful. Such knowledge is obtained through the site characterization process, and is often very difficult to obtain even when possible.

Site Characterization

Site characterization is the process of determining if potential environmental problems exist at a particular site and, if such problems do exist, obtaining enough information to support remedial actions for the site (Melville, 1991; American Society For Testing & Materials, 1996).

The site characterization process (which includes a preview of Response Surface Methodology and its role in site characterization) shown graphically in Figure 1, generally includes the following steps:

1. Defining the objectives of the study and developing boundary conditions.
2. Collecting field data.
3. Developing a conceptual model from the collected field data.
4. Creating a mathematical model from data collected in the conceptual model.
5. Performing a sensitivity analysis on the mathematical model.
6. Validating the mathematical model by performing field experiments.
7. Making future predictions based on the validated mathematical model.
8. Refinement of the conceptual model based on conceptual model tests.

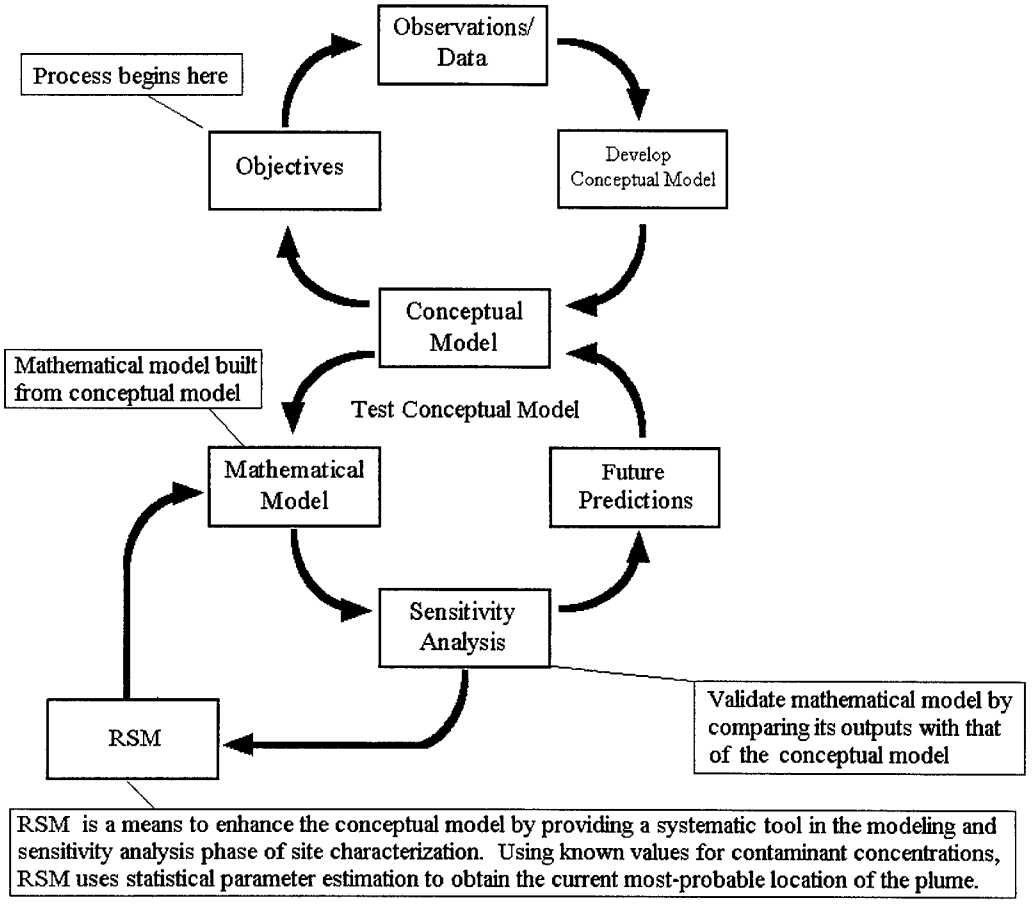


Figure 1. Site Characterization Process and RSM (Adapted from Heiderscheidt, 1996)

Site characterization is also an iterative process, with the number of process iterations of Figure 1 dependent on the initial objectives.

Determining the *objectives* of the site characterization consists of a preliminary analysis of why the site needs to be characterized and what knowledge is hoped to be learned from it. A pre-established set of objectives is critical to the success of the site characterization as it provides clearly defined goals and success conditions for the site. Without such pre-established objectives, for example, a site remediation effort could

continue even where such continued remediation is unnecessary, or be concluded before an acceptable amount of cleanup is attained. These objectives can often be posed as a series of questions such as (Ford and Turin, 1985; ASTM, 1996; Domenico and Schwartz, 1990; Bedient et. al., 1994):

Has contamination occurred?

Where is the contamination located?

What is the source of the contamination?

What are the properties of the contaminant?

What are the site-specific environmental characteristics?

Where is the contaminant likely to go, and how will it get there?

Although these questions are somewhat general in nature, an actual site characterization would include more specific objectives and would ideally result in a more efficient and useful characterization process.

Once objectives are clearly defined, the remainder of the site characterization process always involves the creation of a conceptual model, a representation of the current state of the system. A conceptual model is a clear, physical description of the operation of the system and incorporates all properties of the system that are relevant to the objectives of the study. A well-defined conceptual model also clearly identifies what the system looks like today and the processes that will affect what the system will look like in the future. A well-defined conceptual model also provides an understanding of the

contaminant plume migration within the site's subsurface environment (Franke, 1987; Domenico and Schwartz, 1990; Bedient et. al., 1994; Heiderscheidt, 1996).

One Technique for Enhancing the Conceptual Model

A method currently in use to enhance the conceptual model of a site generally involves a simple mapping of the contaminant plume through the placement of a series of measurement wells. This method, which uses a trial-and-error methodology for well placement based on an analysis of the data obtained from a previous set of trials, continues to be the preferred method for localizing a contaminant source (EPA, 1989). The Environmental Protection Agency (EPA) has identified this method as the *phased approach*, and as the most efficient practical method for the installation of measurement wells. Well locations are determined from groundwater flow directions estimated from existing site data, and measurements are taken. Each subsequent iteration of investigation and measurement requires the distance these wells are placed from a starting point be increased until the approximate area of the plume is located, after which hypotheses and existing information provides the basis for placing additional measurement wells until a map of the contaminant plume is complete. This iterative approach allows the remediation program manager (RPM) to assess the data and revise the conceptual model of the site at each step prior to committing additional resources.

However, this approach produces a large amount of uncertainty, since the ability to ascertain which of these potential source(s) is the actual source through sampling is an

impossible task due to the sample volume obtained from measurement wells being simply too small.

Quantitative Testing of the Conceptual Model

Quantitative testing (QT) provides a formal framework by which to test whether the parameter estimates used in the conceptual model combine to produce the conditions being observed in the field. For example, QT can be used to determine whether a source location is a candidate for a leak given recent contaminant measurements.

To accomplish this quantitative testing of a site's conceptual model, mathematical models (dynamic representations of the system through a series of mathematical equations and procedures) are created using the processes and existing conditions developed in the conceptual model (Franke, 1987; Heiderscheidt, 1996).

Mathematical models often simulate the movement of groundwater and contaminant by computing the movement of the flow of water and contaminant through discrete subdivisions of a site. These subdivisions are typically associated with a cube (Figure 2) representing a portion of the aquifer large enough to represent the properties of the porous medium, yet small enough that any change in the property of interest (i.e., head, concentration, etc.) within this cube is relatively small.

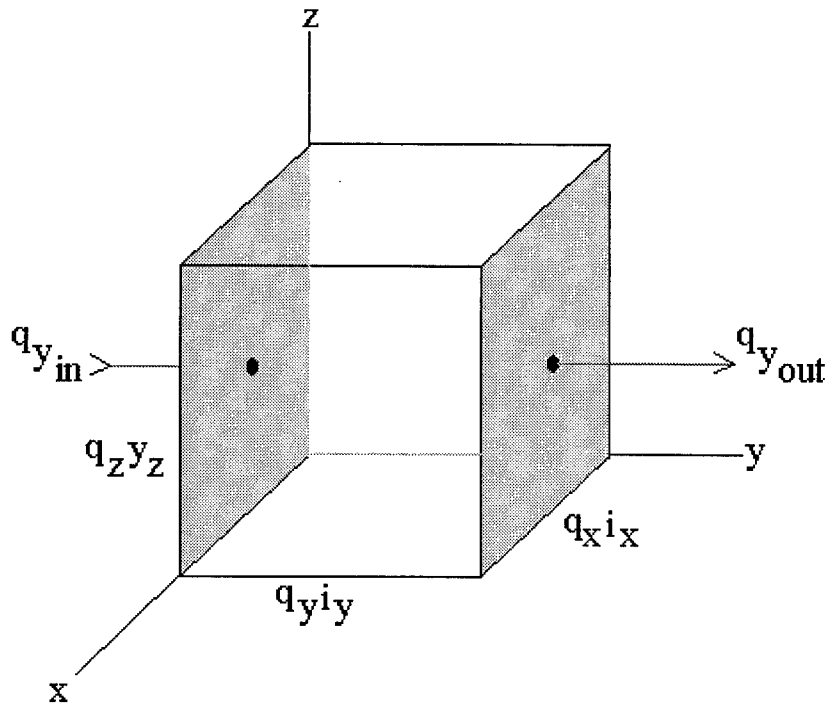


Figure 2. Example of a 3-dimensional finite element (Adapted from Cotman, 1995).

The flow of water through this volume, for example, can be expressed in terms of the specific discharge rate q, which may be written as

$$q = q_x i_x + q_y i_y + q_z i_z, \quad (1)$$

where i_x, i_y, and i_z are the orthogonal unit vectors along the x, y, and z axes respectively, and q is obtained from the three-dimensional Darcy Law equations.

The aggregate of these cubes, which attempt to model an actual site, is often referred to as a grid or mesh. Discretizing the continuous equations of flow and transport on such a grid or mesh is referred to as the finite element method. In this method, the separation between each adjoining subdivision (or element), called a flux boundary, passes information (such as the rate of change in storage or flux, q_{in} and q_{out}) between adjacent

elements in the grid. In this manner, a constructed finite-element model (FEM) discretely simulates the continuous flow of water and contaminant at the site.

Once the FEM is so constructed, hydrologists developing a site's quantitative model have the additional difficult task of assigning correct values to parameters in those models. This process requires correctly assigning parameter values in the models based on information known about the site.

In subsurface hydrology, for example, hydraulic conductivity is often found to be log normally distributed. This type of distribution has the advantage of representing only positive numbers, consistent with the physical requirements of the system. However, it remains difficult-to-obtain correct hydraulic conductivity values for each node in the mathematical model of a site and values for hydraulic conductivity are therefore generally calculated from other known parameters (Gelhar, 1993).

As a result, mathematical models used in conjunction with sensitivity analyses often serve as efficient tools for calculating these difficult to obtain parameters. Such models incorporate a set of processes which is equivalent to *solving the inverse problem* (Carrera, 1987).

The Inverse Problem

The inverse problem can be more simply described as a series of known and unknown parameter values in a system. The known values are those which, in a system of hydrodynamic equations, are usually the dependent variables. The unknown values are the

independent input variables. An example of this relationship can be found in the logical equation,

TIME + SPACE + INITIAL CONCENTRATION = CONCENTRATION, (2)

where CONCENTRATION and TIME are known and INITIAL CONCENTRATION and SPACE are unknown.

The difficulty with estimating the parameters for INITIAL CONCENTRATION and SPACE is that often not enough information exists to solve for them uniquely. Such a problem is *ill-posed* since an infinity of solutions exist for both SPACE and INITIAL CONCENTRATION.

Ill-posed Inverse Problem and Migrating Contaminants

Understanding contaminant migration in the saturated zone requires quantitative representations of advection, dispersion, and other processes applicable to the affected site. (Mackay, 1985). Once these quantities have been estimated and a model constructed, locating the contaminant plume becomes a problem of identifying the contaminant source values (for location and contaminant mass flux) which would produce the measured data in a numerical model of the site. Obtaining these source conditions is thus equivalent to that of finding a solution to the inverse problem.

Unfortunately, locating a migrated contaminant source is in almost every case an ill-posed inverse problem, as it is very difficult to find exact values for the input parameters (source location and mass flux concentration) of the migrated contaminant point source. Note that although in reality a contaminant source is not a point source, modeling migrated contaminant sources as a single point injection is commonly done in

current site remediation technology, which typically yields good predictions for the movement of a contaminant plume (Freeberg, 1987). Point sources are used in quantitative models since the locations (nodes) represented in the model represent the typical value for a particular volume around the node. That is, the quantitative model consists of the discrete collection of data points. This work employs a finite element grid with nodes at 100 foot intervals.

Since several parameters can be adjusted to yield the same measured contaminant levels, a multitude of different locations and contaminant mass flux values will yield the same measured concentration levels (Cotman, 1995). Such a system is ill-posed primarily because of the limited data that can be feasibly collected from sampling or observation wells. Stated differently, the relatively small number of wells which are currently implemented at remediation sites do not provide enough information to model the site in a well-posed manner, and obtaining enough sampling data to make the problem well-posed is prohibitive with respect to cost.

Response Surface Methodology

One mathematical tool little used in site characterization is Response Surface Methodology (RSM), an indirect approach to solving the inverse problem. RSM is categorized as an indirect method because input parameters are estimated and model solutions are compared with actual measured values, as opposed to a direct method which

treats unknown parameters, for example, hydraulic conductivity and storativity, as dependent variables in a formal inverse boundary value problem (Khan, 1986).

RSM uses statistical techniques for empirical model building and exploitation which allows for finding the best fit of a number of input variables to a model's output. Through systematic and simultaneous adjustment of multiple hydrogeological parameters, and through analysis of both measured values and model outputs, the technique yields good values for the source conditions even in ill-posed scenarios. For this reason RSM can prove ideal as a systematic source condition or parameter estimation tool in site remediation (Cotman, 1995).

USGS SUTRA Numerical Model

RSM has the advantage that it may be used with virtually any environmental flow and transport model (e.g., SUTRA, MODFLOW/MT3D, etc.) For the work presented here, the USGS SUTRA (Saturated-Unsaturated TRAnsport) model was chosen. SUTRA is a software modeling program that simulates fluid movement and the transport of dissolved substances in a subsurface environment, employs a two-dimensional hybrid finite-element and integrated-finite-difference method to approximate the governing equations, and provides as the primary calculated results fluid pressures and solute concentrations as they vary in time everywhere in a simulated subsurface system (Voss, 1984).

Summary and Research Goals

Scientific processes which result in accurate estimates for the location and magnitude of a contaminant source have great potential for reducing the overall cost of site characterization and remediation. Quantitative modeling gives us a formal process by which to test conceptual models. RSM as a mathematical tool provides a systematic approach to the formal process of estimating source parameters through an indirect solution of the inverse problem.

This work will primarily demonstrate the effectiveness of RSM as a source condition or parameter estimator in the site characterization process. The use of RSM and the methods outlined in this work have great potential for supplementing site remediation programs worldwide by localizing the current location of a migrated contaminant source.

III. Theory and Methodology

Introduction

This section outlines both the steps needed to successfully implement RSM and the theory behind how RSM can be applied to the groundwater and contaminant transport problem.

Response Surface Methodology and Screening Theory

RSM uses statistical techniques for finding the best fit of a number of input variables to a model's output. By systematic and simultaneous adjustment of source locations and contaminant mass flux values, and through an analysis of both measured values and model outputs, RSM provides a solution set of locations and contaminant mass flux values of continuous sources (Box and Draper, 1987).

The particular steps needed to characterize a site and obtain solutions of the contaminant migration inverse problem (previously outlined in detail in Figure 1) which are relative to using RSM include *creating a mathematical model based on the conceptual model*, *validating the mathematical model*, and *updating the conceptual model*. RSM can accomplish these steps by:

1. Assembling the measured contaminant data. This data will be used with model run values for the same points (nodes) in the mathematical model to calculate a sum of squares error (SSE). The SSE is defined as the sum of differences between the model output and actual values for given locations

2. Estimating the amount of time the contaminant leak has been leaking.

3. Identifying an initial range of contaminant flux values.

4. Identifying locations on the site which used chemicals of the type found in the measurement wells.

5. Establishing regions of possibility and interest based on the locations identified in step four above and through an analysis of contaminant levels. This process is described in detail later in this chapter.

6. Identifying nodes which will be screened within an area of interest.

7. Using Plackett-Burman (PB) designs to minimize the number of model runs based on the number of nodes within the area(s) of interest identified in 6. above.

8. Screening the identified nodes using the selected PB design to obtain the one or two (or group of) significant nodes.

9. Performing additional screenings where necessary to isolate a single node from a group of significant nodes (see 8. above).

10. Using RSM methods to obtain the mathematical model's input flux value for the significant node(s) identified in 8 or 9 above. This step requires adjusting the node's flux value between three (upper, lower, and median) values and observing the value of the SSE as the input flux value changes. The input flux value which minimizes the SSE is the solution value for the selected significant node.

11. Updating the conceptual model as necessary.

Setting up RSM to locate a migrated contaminant source is also a two-part process. Part one of the process essentially identifies the most significant locations for the migrated contaminant source and part two determines the flux values for that selected source.

Part one of the two-stage process includes screening techniques that determine which model locations significantly influence the value of the SSE responses. Although there are a variety of ways error can be measured, in this work the SSE is defined as:

$$SSE = \sum_{i=1}^{n} (h_m - h_s)_i^2 , \tag{3}$$

where h_m represents the current measured contamination values at n chosen nodes at the remediation site and h_s is the model output concentrations for each of those n modeled nodes or locations.

Part two of the RSM process uses a first-order (linear) approximation of the SSE response surface to identify local optimal parameter settings that minimize the SSE for the model. The response surface is defined as the n-dimensional surface obtained when plotting the SSE for each possible parameter value.

A key feature of RSM is the empirical testing of the response surface. Instead of determining the shape of the entire response surface, linear approximations or *patches* of the surface are approximated and adjusted by tests in the direction of suggested improvement. This empirical testing, which leads to an understanding of the underlying SSE response without the need to precisely define it, is a key advantage of RSM over most other inverse methods (Box and Draper, 1987.)

Part I - Factorial Designs and Parameter Screening

The first step in locating a migrated contaminant source using RSM is to perform a series of screening experiments to determine which potential source locations are significant contributors to the error surface. This screening serves the same purpose as a sensitivity analysis in traditional groundwater modeling studies. However, determining which locations are significant is a difficult process that requires human interaction with the data and special knowledge of the groundwater system being studied (Box and Draper, 1987).

Factorials are statistical designs which greatly facilitate the process of making comparisons, seeking similarities, and noting differences and trends with the groundwater system, and more specifically, the specific locations under scrutiny in the model. These designs possess many desirable properties, to include a) facilitating model creation and criticism by allowing many comparisons to be made at once, b) providing extremely efficient estimates of parameters, and c) providing for simple calculations. (Box and Draper, 1987).

Complete factorial designs for *k* factors (*k* model locations) is obtained by choosing N_1 discrete *levels* of factor 1, N_2 *levels* of factor 2,..., N_k *levels* of factor k, and then executing the $N = N_1 \cdot N_2 \cdot N_3 \cdot ... \cdot N_k$ model runs obtained, which provide all possible combinations of the *levels* selected (Box and Draper, 1987). In this work, the *levels* for each of the k factors or parameters are binary in nature and are obtained from the preselected high and low contaminant mass flux values. That is, parameters are set to the

high and low levels of the *expected* mass flux range based on observed local hydrogeological conditions and historical information about the types of contaminants used at the site. Note that RSM will still work well even if actual values are outside this range. These parameters are later coded into standardized (logical) variables, resulting in $N_1 = N_2 = N_3 = ... = N_k = 2$ levels for each factor. Examining 19 locations (typically used in this work) would therefore normally require 2^{19} (nearly one million) model runs using a full factorial design. This large number of model runs essentially make using full factorial design unfeasible.

Plackett-Burman Designs

However, Plackett-Burman (PB) designs are first and second order designs that allow for estimation of the k main effects (factors) in only $k + 1$ runs. This is a significant reduction from a full factorial design of 2^k runs. Since this work employs first order approximations (mainly for simplicity of calculations) and since the main effects of each location can be obtained with a first order design, the use of PB designs is ideal for reducing the working set of potential parameters (Box and Draper, 1987).

Almost any number of variables can be represented in a PB design. An example of such a design for 19 variables is obtained from a cyclical permutation of the logical sequence 1 1 -1 -1 1 1 1 1 -1 1 -1 1 -1 -1 -1 -1 1 1 -1 (where 1 represents the high level of the expected range and -1 represents the low level). The permutating sequence continues by taking the logical sequence from the first (previous) model run and rotating the last

digit to the first position for the next rows, with the last (20th) row in the PB design consisting of all -1's. A completed 19-factor PB design is shown in appendix B.

Once the model runs have been performed, the SSE responses obtained from the 20 runs are fit to a first-degree polynomial in 19 coded variables (Cornell, 1990) of the form

$$Y_u = \beta_1 x_{u1} + \beta_2 x_{u2} + \ldots + \beta_{19} x_{u19} \qquad (4)$$

where Y_u is the SSE response obtained from the run, β is the coefficient of the variable X, and u is the particular run of data. Since only those locations with significant first-order effects need be considered in part two of the RSM process, a normal probability plot can be used to determine which of these locations are significant. This plot is obtained as follows:

1) *Effects* obtained from the least squares solution are ordered from smallest to largest, with the x-axis scaled accordingly. The *effect* for each of the *k* locations used in the PB runs and defined as the amount of significance of each location, is calculated from the coefficients of the least squares solution of the $k + 1$ equations and k variables (locations). In this 19-coefficient, 20-run model example, the *effect* is obtained from the coefficient values of X_1, X_2, X_3, ... X_{19}, and is defined as

$$\text{effect}_k = 2 \cdot \beta_k \qquad (5)$$

2) The quantity ϕ^{-1} [(i - 0.5)/k] is plotted, where ϕ^{-1} (p) is the inverse cumulative distribution function of the standard normal distribution and *i* represents the rank (an ordering from low to high) of each *effect*. This function can be approximated by equation (6) (Box and Draper, 1987).

$$\phi^{-1}(p) = [p^{0.1349} - (1-p)^{0.1349}] / 0.1975 \quad (6)$$

3) Modeled source locations which significantly contribute to the output of the groundwater model are selected for further analysis. In a properly fit first-order linear model, the residuals (ordered *effects*) are approximately normally distributed and have equal variance. If there are no significant *effects*, each of the computed *effects* represents an observation from a common normal error distribution. A plot of these *effects* on the normal probability paper of the empirical cumulative distribution function should roughly form a straight line. However, if some of the plotted points differ noticeably from a straight line either at the top right or bottom left, then the corresponding *effects* are considered *significant* (Cotman, 1995). Figure 3 shows an example of two locations (circled) which have been selected as significant.

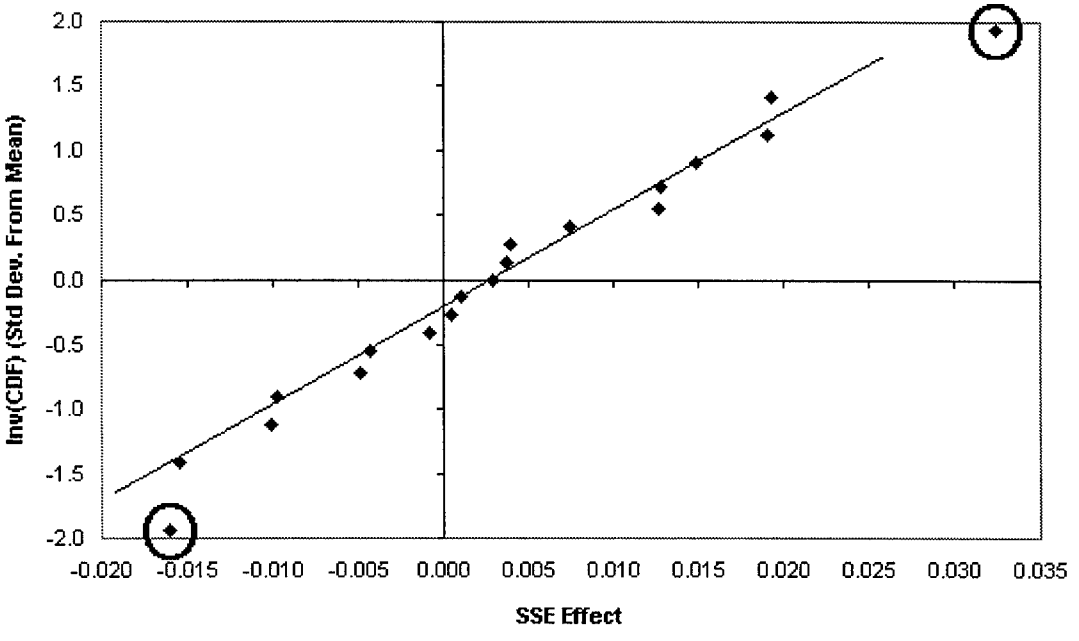


Figure 3. Significant locations on a sample SSE normal probability plot.

In part two of the RSM process, the settings of these parameters are adjusted to reduce the value of the SSE response to within a desired tolerance.

Part II - First Order (Linear) Designs

RSM investigates the nature of the SSE response surface by conducting designed experiments centered at the significant source locations identified in part one of the process. A two-level, full-factorial (2^k) design is needed to empirically evaluate the SSE response surface. In the sample SSE plot shown in Figure 3, the two selected *significant* model locations require $2^2 = 4$ runs, as opposed to the 2^{19} or nearly one million runs that would be needed if each of the 19 example locations were considered significant possibilities for the contaminant source. This reduction illustrates the importance of the screening phase since the number of model runs grows exponentially with the number of significant nodes identified.

Using the responses obtained from individual model runs, an N-dimensional response surface is generated (for a model with N-1 significant effects). Figure 4 illustrates the three-dimensional response surface for the two-parameter system of Figure 3. This surface is obtained by plotting values of the two parameter inputs (representing the X and Y axes in a three-coordinate system) and the resulting SSE for those parameters (representing the Z axis).

The purpose of part two of the RSM process is to find the values for each parameter that will cause the SSE to decrease the quickest, that is, a coasting or sliding in the direction of steepest descent. Multiples of these steepest descent parameter values,

referred to in this work as the *reduction gradient*, are used in a number of model runs to determine the parameter value(s) which minimizes the SSE.

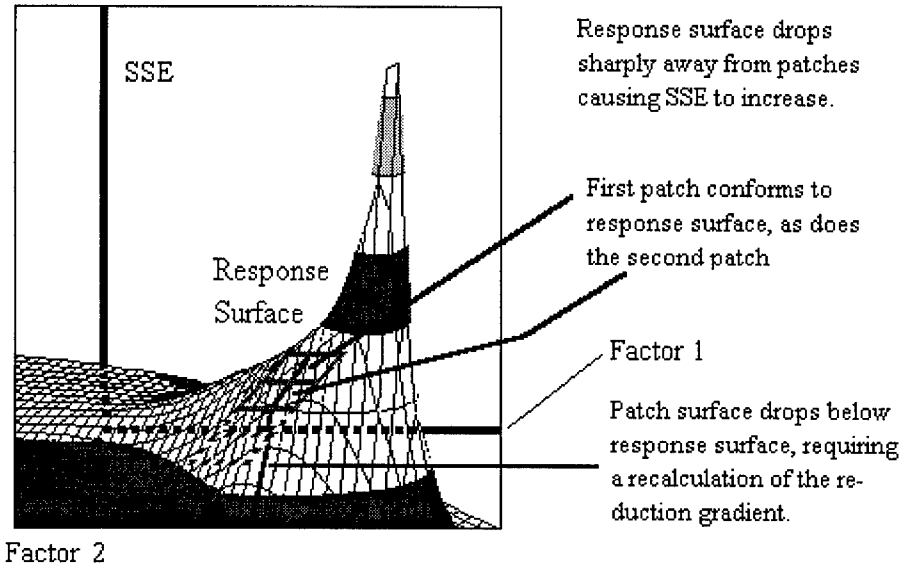


Figure 4. Reduction gradient on a 3-D response surface.

Two steps are required to obtain the reduction gradient. First, a linear approximation or patch of the response surface is developed. Second, the reduction gradient of the patch (a trivial matter for a linear approximation) must be calculated. The center of each patch, obtained from the original parameter values, lies on the response surface. Each model run then requires this patch be extended in the direction of steepest descent. Once the SSE begins to rise along this reduction gradient, the reduction gradient is no longer valid. When this occurs, as shown in Figure 4, a new linear patch must be formed and the reduction gradient must be recalculated at the point of lowest SSE along the previous gradient. Additional model runs along the new reduction gradient is then accomplished.

As the reduction gradient extends the patch, SSE responses will typically improve (decrease the SSE) for one or more runs and then worsen as the model diverges from the underlying response surface, as shown three-dimensionally in Figure 4 and two-dimensionally in Figure 5.

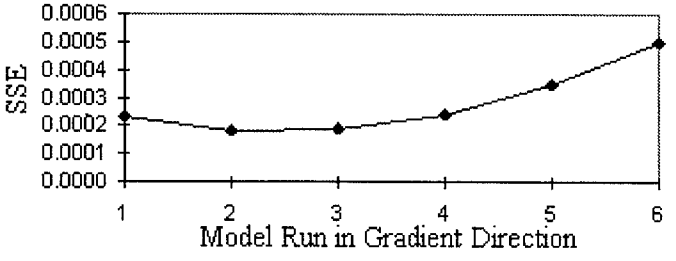


Figure 5. Linear model and divergence from SSE surface for consecutive model runs.

If the SSE value at its lowest point on the reduction gradient is acceptable (within a certain predetermined tolerance) or if no significant changes occur as a result of a newly calculated reduction gradient (such as is found when the response surface forms a flattened local area or a shallow bowl,) the local minima of the response surface has in effect been located. The parameter values last used are those which are statistically the best values for those locations chosen from part I, and yield the least error between the contaminant levels produced by the selected locations and the *measured concentration value*.

Fuzzy Logic Theory

Basic principles of fuzzy logic will be applied to some elements of this work, mainly to provide a method for quantifying the amount of uncertainty a node or location has for being the source of a migrated contaminant.

With the ultimate aim of discovering better methods for modeling the real world in mathematical terms, an increasing number of mathematicians and computer scientists are turning to fuzzy logic and fuzzy systems (the collection of fuzzy logic variables together with their associated operations). Fuzzy logic, an alternative to the traditional Aristotelian methods of logic, developed by L. A. Zadeh in 1965, includes a modification of both the standard definitions of set membership and the single state (one or zero, on or off, etc.) of a proposition's *truth.* Zadeh describes fuzzy logic as a system to provide a model for modes of reasoning that are approximate rather than exact, and derives the importance of fuzzy logic from the fact that all of human reasoning--and especially common-sense reasoning--is approximate in nature (Zadeh, 1990).

A key concept to fuzzy logic is that binary truth values are no longer restricted to zero and one. Instead, a variable can take on truth *membership values* anywhere in the continuous range from 0 to 1, with membership values obtained from measured data (Kandel, 1986). For example, if collected contaminant measurements provide rough concentration data between 0.0003 and 75 parts per million (ppm) of compound X, those locations with concentrations near 75 ppm could be assigned truth values of *Contaminated*

near 1.0, while those sites with concentrations near 0.0003 ppm could be assigned truth values closer to zero, with the region between the two values scaled exponentially.

Applying Fuzzy Logic to Regions of Possibility

In this work, the concepts of fuzzy logic, hydrogeological data, and concentration levels were combined to generate cone-shaped *regions of possibility*, that is, possible locations for the migrated contaminant point source. Each cone is generated with its height controlled by the maximum a contaminant source could have migrated in the period identified in the model. The tip of the cone is placed on the location of the initial (historical) leak or spill. The direction of the cone's bisecting line is placed in the direction of the gradient or groundwater flow direction, but could also be modified to incorporate known site data such as the general dip and strike of the soil near the original leak or spill around which the *region of possibility* is being drawn (Poulsen and Kueper, 1992). The cone's base is roughly generated by factors such as the amount of dispersivity in the soil.

The conical region can be seen as a type of fuzzy structure, since the values for the migration distance, dispersivity, groundwater flow rate, etc. are uncertain. This structure can then be developed further by examining *high possibility* contaminant levels in the immediate area. These levels were established by identifying the maximum contaminant concentration found in a region. The *source possibility* of a particular well being the contaminant source was then defined as the percentage its measured contaminant level is of the maximum contaminant level. This work generally focused on those wells with a

source possibility of 0.7 or greater. The conical region of possibility is then enlarged or reduced to bound the wells with the predetermined source possibility.

Areas of Interest and Bounding

Areas of interest are then developed to place special emphasis on those wells that measured exceptionally high concentrations, as migrated contaminant sources are likely to be near these wells. These areas of interest are identified by wells with lower source possibilities on the perimeter of the region and higher source possibilities in the center of the region. The areas of interest are ideal for use in RSM calculations, which produce a candidate or candidates within the area for the current location of a migrated contaminant source.

If areas of interest cannot be generated because of an insufficient amount of data (wells), additional measuring wells can be sunk until a) these regions are formed or b) the region of possibility is eliminated as a candidate for the point source.

Special care must be taken if a particular area of interest is identified but with very low *source possibility* (low concentration comparatively) internal wells, as this may identify a region that contains a contaminant source separate from the higher source yet with a lower total mass flux. This type of area of interest, which can be difficult to find because of the overwhelming nature of the larger contaminant source, should be evaluated with the maximum *source possibility* that is local to the region.

Where several contaminant sources are indicated, areas of interest can be generated as previously described. Where the potential for overlapping contaminant levels is high,

areas of interest are screened together. Where areas of interest are far enough apart so that there is minimal concentrate mixing near measurement wells (regions are independent), each area of interest can be screened and processed separately.

Summary

Given the screening theory, factorial designs, first order statistical designs, and fuzzy logic theory, a foundation for building methods for locating migrated source contaminants can be employed in a set of three experiments. The next section will outline how these elements were used together to generate results.

IV. Experimental Data

Introduction

Since the emphasis of this work was to evaluate how well migrated contaminant sources can be located statistically using RSM, a numerical model based loosely on the Hill AFB Operable Unit 3 site was constructed (Montgomery-Watson, 1995). This site provided a generic model from which to work -- one which reflects a realistic flow and transport scenario where source uncertainties could have significant impact. However, the models used in this work do not represent actual contaminant conditions at Hill AFB.

This work demonstrates the usefulness of the RSM process on this numerical model through three separate experiments. In these experiments, several factors were evaluated, including the effectiveness of RSM given a) a single unknown source with an unknown flux value, b) multiple unknown source locations and multiple, unequal flux values, and c) differences between the experimental model (set up with nodes having average estimated values for hydraulic conductivity) and a modification of model nodes through the use of log-normally distributed hydraulic conductivity values on the order of the Borden Site (Sudicky, 1986). Although these variations should provide enough real data to determine the effectiveness of RSM techniques, practical demonstration of RSM techniques on an actual contaminated site is left for future work. Clearly, when using the techniques outlined in this work on an actual site, the groundwater flow / transport numerical model must be designed as close to the site's actual hydrological and geological

conditions as possible for RSM to be effective, although the third experiment in this work demonstrated that an exact model is not a requirement for obtaining usable solutions.

Hill Air Force Base Operable Unit 3 and Modeling the Problem

Hill Air Force Base is located in northern Utah, approximately 25 miles north of Salt Lake City. It occupies nearly 7000 acres in Davis and Weber counties, and is the location of an intense remediation effort by the U.S. Air Force. Several operable units (OUs) are being studied on Hill Air Force Base, including Operable Unit 3 (OU3), located near the base's southern boundary.

OU3 is located in the heavily industrialized portion of the base. Seven potential source areas have been identified, several of which are associated with the present and past treatment of industrial wastewater. Industrial activities using hazardous materials, storage tanks, and pipelines are also prominent in this area.

OU3 consists of a variety of soil layers to include sand, silty or clayey sand, silt, sandy clay and silt, and clay and silty clay of the Provo Formation. This site has also been well-characterized in the Phase II Remedial Investigation Report for Operable Unit 3 (Montgomery-Watson, 1995).

The *measured concentration* or truth *values* for the migrated contaminant in the modeled site in this work is obtained from several runs of the numerical model based on the Hill AFB OU3 site (hereafter referred to as the alpha model) with several locations (nodes) chosen by an independent party to represent the migrated contaminant sources.

The contaminant sources chosen were kept separate from the experiment until the completion of the RSM experiment. The concentration map obtained from the model runs were then used as the truth sets from which actual concentration levels were obtained. For example, as additional wells were placed on the alpha site, concentration values from the truth set were used to simulate the actual drilling of wells. Figure 6 identifies the concentration values obtained in the truth set run for experiments one and two, and outlines two sources of contaminants: a high concentration source located near the top right part of the grid and a lower concentration multiple-source near the bottom right portion of the grid. Potential (historical) sources of leaks are shown in the context of existing structures and locations in Figure 7, based on the Phase II Remedial Investigation Report for Operable Unit 3 (Montgomery-Watson, 1995) as buildings, ponds and other facilities on Hill Air Force Base in the area of OU3. These were included in the alpha model in generally the same relative locations.

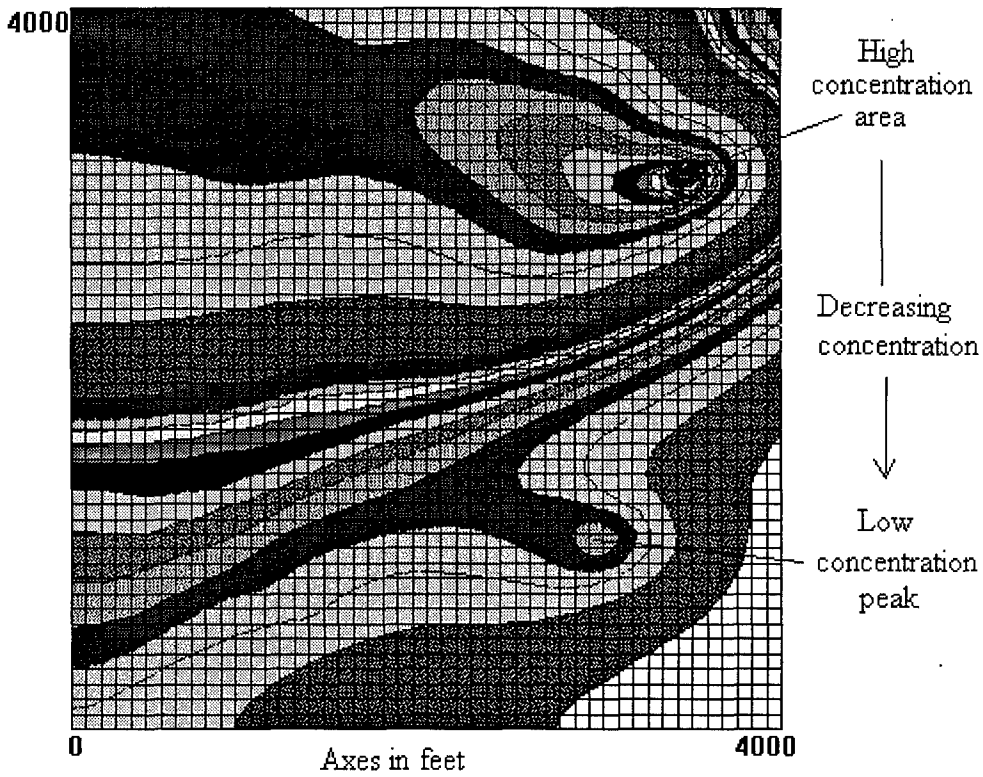


Figure 6. Contaminant levels in truth set for experiments 1 and 2.

Numerically modeling the alpha model contaminant transport with groundwater flow from known geological data (such as head values, soil content, etc.) produced the hydraulic head contours shown in Figure 7. These contours were generated by adjusting the pressure differences at various key grid points on the finite element grid and estimating both the pressure difference between the east and west boundaries of the site and the permeability of the system.

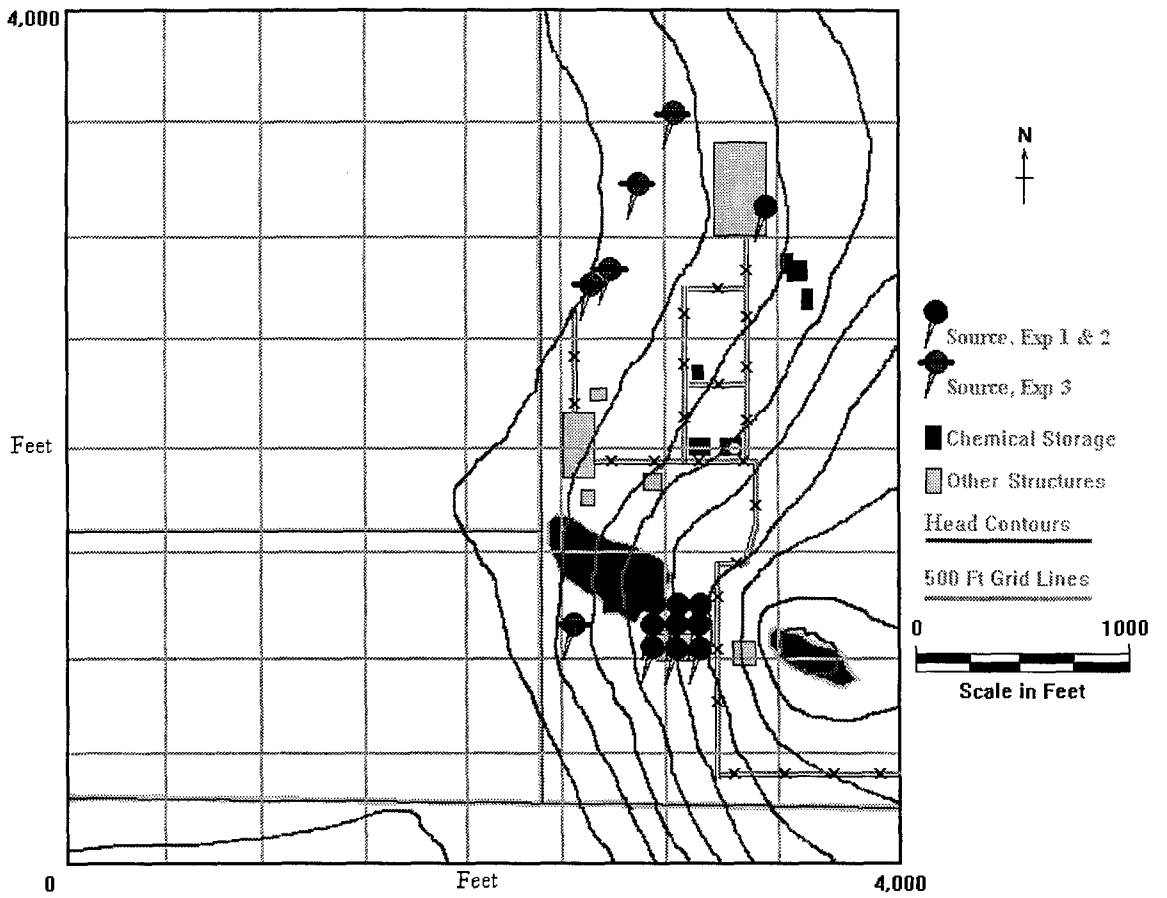


Figure 7. Groundwater flow lines and modeled locations of contaminant sources.

In experiment three, flow lines were further modified by employing log normally distributed values for hydraulic conductivity. This experiment used these log normally distributed values to generate flow field and truth concentration values which were different from those used in experiments one and two. However, the same averaged or mean value for the flow field as in experiments one and two was used for the analysis. This experiment was designed to test whether RSM would be effective when the flow and transport models were an imperfect representation of the truth set.

In experiments one and two, two contaminant sources were simulated to test the RSM procedure. (These simulated sources do not represent the actual contamination on Hill AFB.) The first contaminant source, located at grid location (3400,3000), consisted of a mass flux concentration obtained from a programmed relative concentration of 1.0 and flow rate of 4.4×10^{-4} m^3/sec. The second source, centered at grid location (3000,1000), formed a square of source nodes with 200 foot sides. The mass flux concentration for each of the nine locations in this second source was obtained from a programmed relative concentration of 1.0 and flow rate of 1.7×10^{-6} m^3/sec. The second source also modeled the spreading of contaminants by rainwater infiltration, and was used primarily to test the usefulness and accuracy of this work when multiple sources are present. In experiment three (included on Figure 7), the first contaminant source, located at grid location (2500,1000), consisted of a mass flux concentration obtained from a programmed relative concentration of 1.0 and flow rate of 1.5×10^{-4} m^3/sec. The second source formed a line of four leaks which included grid locations (2600,2600), (2700,2700), (2800,3100), and (3000,3400). The mass flux concentration for each of the four locations in this second source was obtained from a programmed relative concentration of 1.0 and flow rate of 1.2×10^{-6} m^3/sec. The second source attempted to model a contaminant source leak along a long, damaged pipe, and was also used primarily to test the usefulness and accuracy of this work when multiple sources are present. At the conclusion of the experiment, the exact locations and mass flux values used in both truth sets was compared with those obtained using RSM techniques.

Estimating the chronology of the leak or spill is also a necessary step for setting up the problem. As a first step to understanding the capabilities of the RSM to find migrated contaminant sources, the contaminant was modeled a) as having migrated to their new source locations in a negligible amount of time, and b) with the same mass flux output of at their fixed migrated locations for the duration of the experiment. A fixed value of 30 years was used as the amount of time the new location has been sourcing contaminants, corresponding roughly to the number of years locations with potential leaks were in operation at the OU3 location.

Finally, a range of values for contaminant flux for each potential source was needed for the screening phase of RSM. In this work, contaminant mass flux was considered as a whole, calculating it using a relative concentration of 1.0 kg/m^3 and a flow rate range of between 0.001 and 0.00000001 m^3/day. These values generated a mass flow rate, which when averaged over the inter-nodal area of the numerical model, produces mass flux values representative of leaks or pipes and tanks, although higher or lower values can be used as needed.

Experiment One

Introduction

In experiment one, the assumption was the site was contaminated yet no wells were placed. This assumption permitted a use of the methods incorporating fuzzy regions developed earlier.

Selecting Candidate Regions for the Migrated Contaminant

Five locations were selected as candidates for which a contaminant source developed and migrated. In the alpha model, we assumed that these five locations were the only possible sources of contamination.

Eight wells were then placed in conical *regions of possibility* around each of these five locations. These *regions of possibility* are shown in Figure 8, and are developed from the ideas presented by fuzzy logic described earlier.

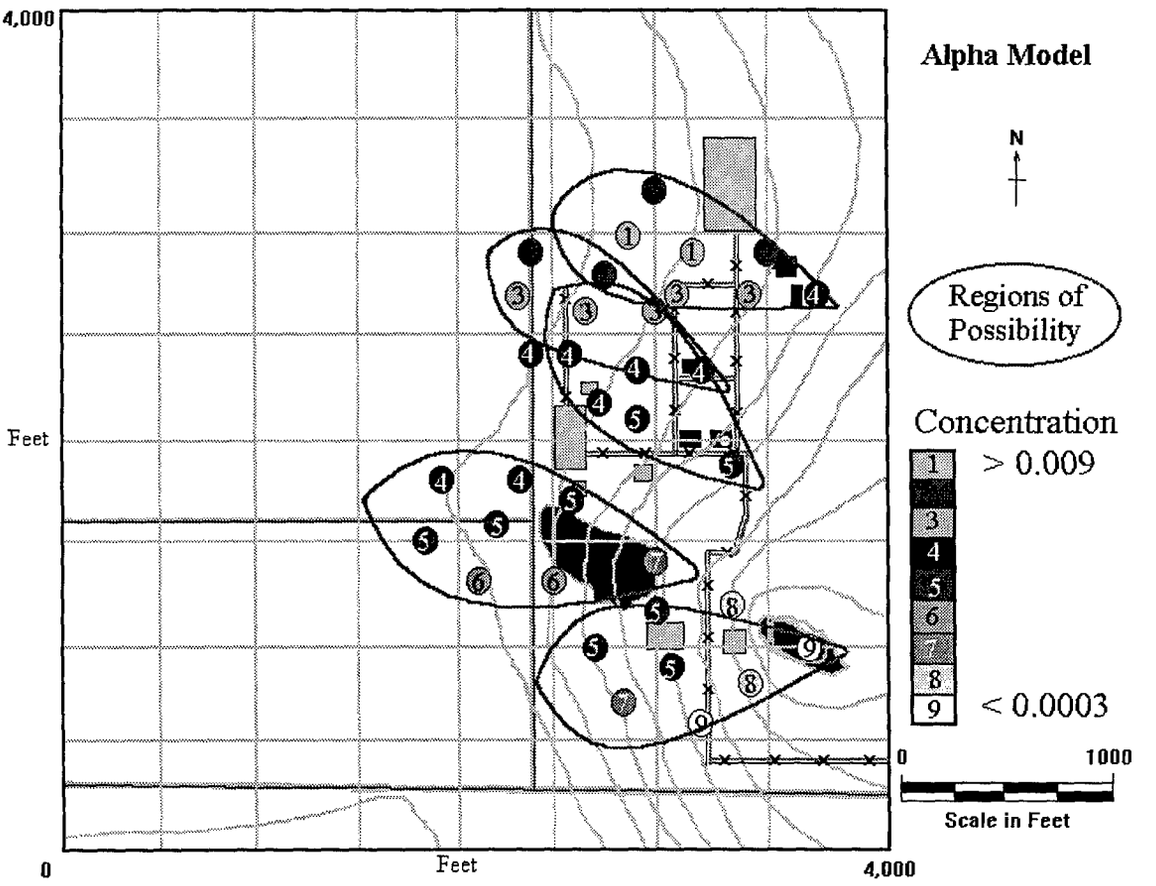


Figure 8. Experiment #1 well locations and concentrations using fuzzy logic.

The measured concentration values of the initial wells located in the five regions then served as the *measured concentration value* from which a sum of squares error (SSE) was calculated (SSE).

After recording and graphing well locations and concentration, each *region of possibility* was evaluated against the criteria for becoming an area of interest, as described in the previous section *Applying Fuzzy Logic to Regions of Possibility*. These selected areas of interest were then bounded, as shown in Figure 9, and a representative number of locations in each region were then used in the first of two screenings.

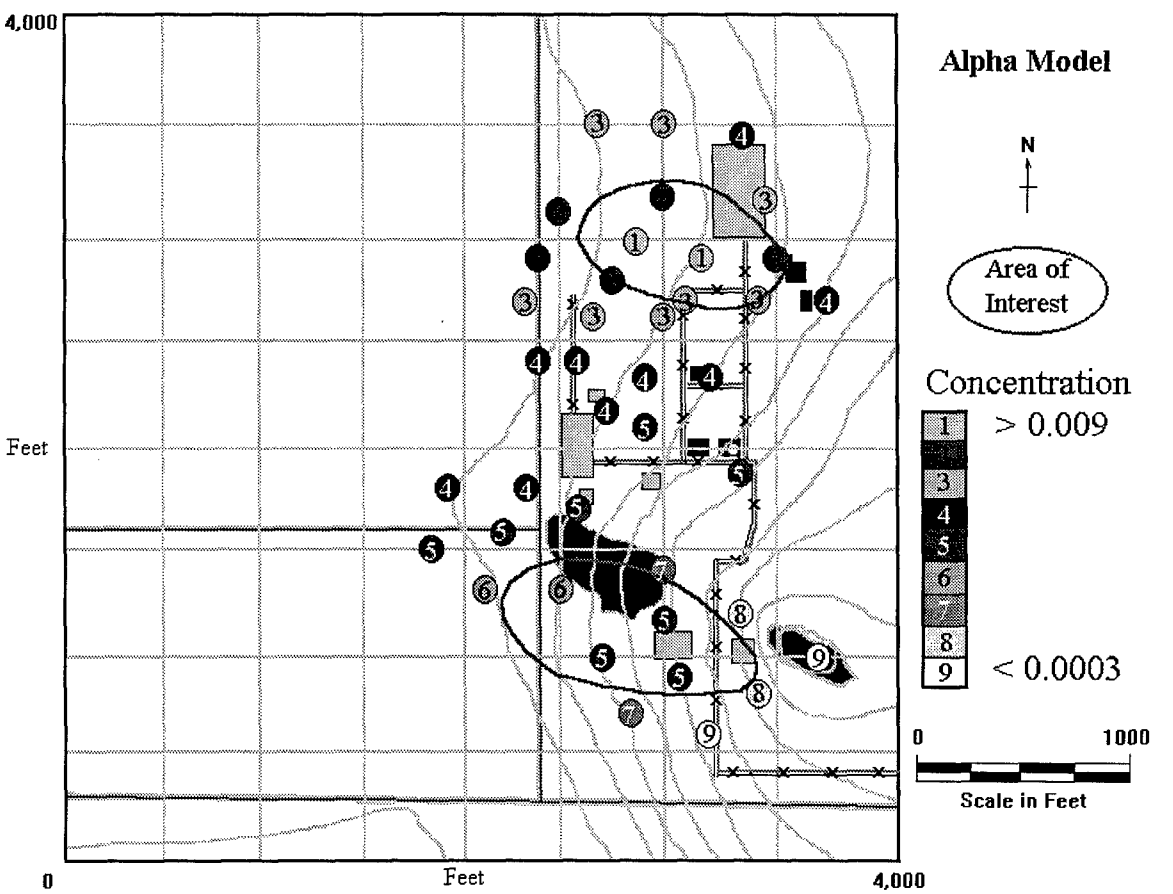


Figure 9. Experiment #1 bounding wells and areas of interest.

Before the screening could begin, several determinations needed to be made. First, since the two *areas of interest* were in relatively close proximity, they were screened together since increasing or decreasing parameters in one area affected model concentration values in the other area. Second, since the lower area of interest had measured concentration values an order of magnitude lower than the upper region, the maximum flux for the lower region was adjusted to an order of magnitude smaller than the level for the upper region.

Screening for Significant Locations

In the first step of the RSM procedure, 101 locations (representing the numerical model locations within the upper and lower areas of interest as shown in Figure 10) were examined. Although not impossible, the large number of locations were too numerous for a single screening run, therefore a smaller number of wells within each of the two *areas of interest* were selected to localize the contaminant. Since the distances between each of these smaller number of wells was larger than the distance between adjacent locations (nodes) in the numerical model (see Figure 11), a second screening was later performed to more precisely locate the source using those nodes not selected in the first screening.

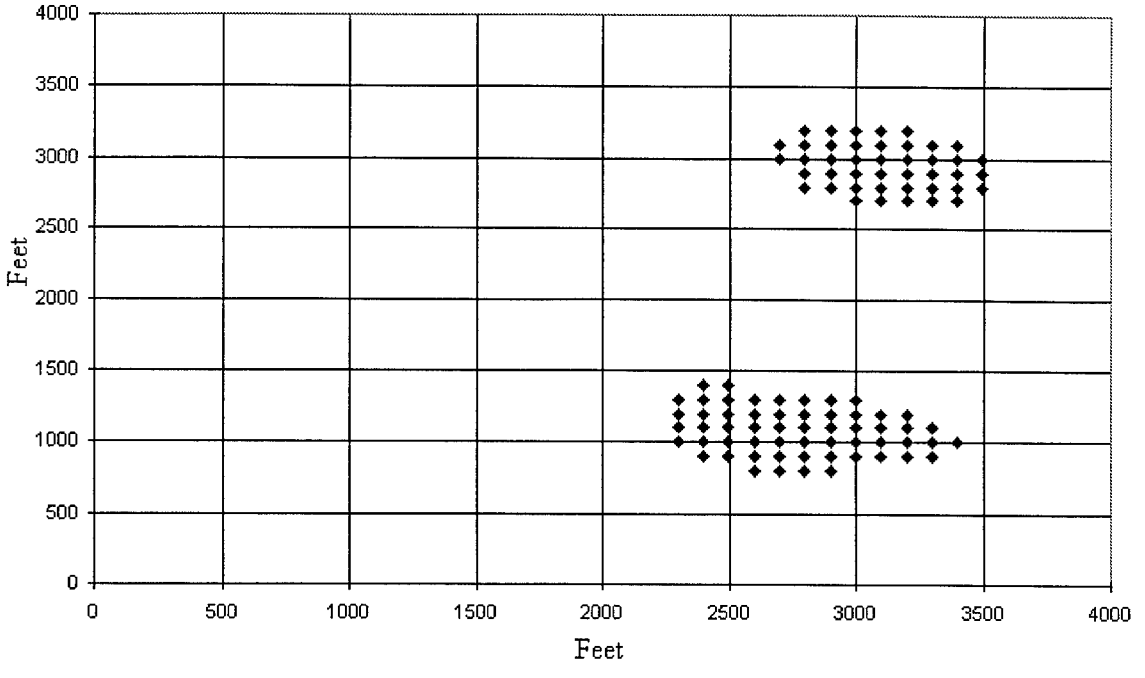


Figure 10. Experiment #1 candidate nodes for initial screening process.

A Plackett-Burman (PB) design for 19 locations (20 data runs) was used in the first screening; 11 locations in the upper site and 8 locations in the lower site. These 19 locations, spread out evenly throughout the two *areas of interest* as shown in Figure 11, consisted of only a small fraction of the 101 locations in the areas of interest and an even smaller fraction of the 1600 locations represented in the numerical model, a scenario that often takes place in the real world.

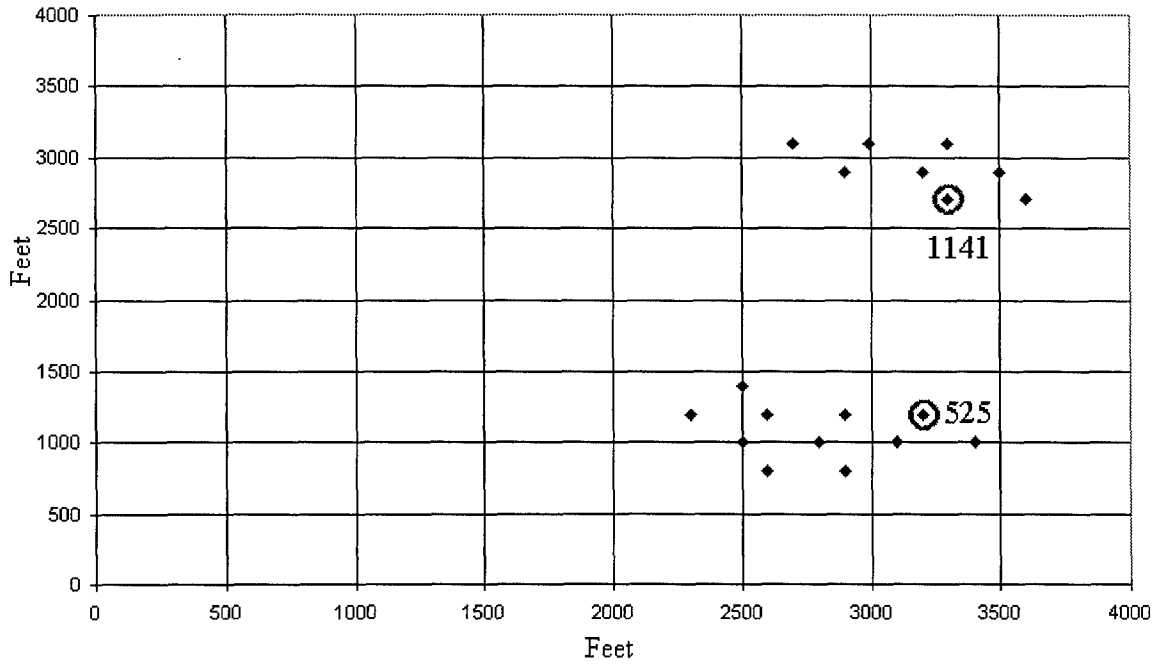


Figure 11. Experiment #1 locations used in initial screening.

The 20 required model runs were completed with each of the 19 selected locations programmed as contaminant sources whose *level* values were calculated using either a minimum (0.00000001 m^3/day) or maximum (0.001 m^3/day for the upper site, 0.0001 m^3/day for the lower site) flow rate.

The concentrations resulting from each of these 20 screening runs were then compared to the *measured concentration values* for determining which of the 19 screened locations were statistically significant, using the normal probability plotting procedure described earlier, as shown in Figure 12. Circled nodes 524 and 1141, corresponding to nodes in the SUTRA model, were identified as significant in this set of model runs.

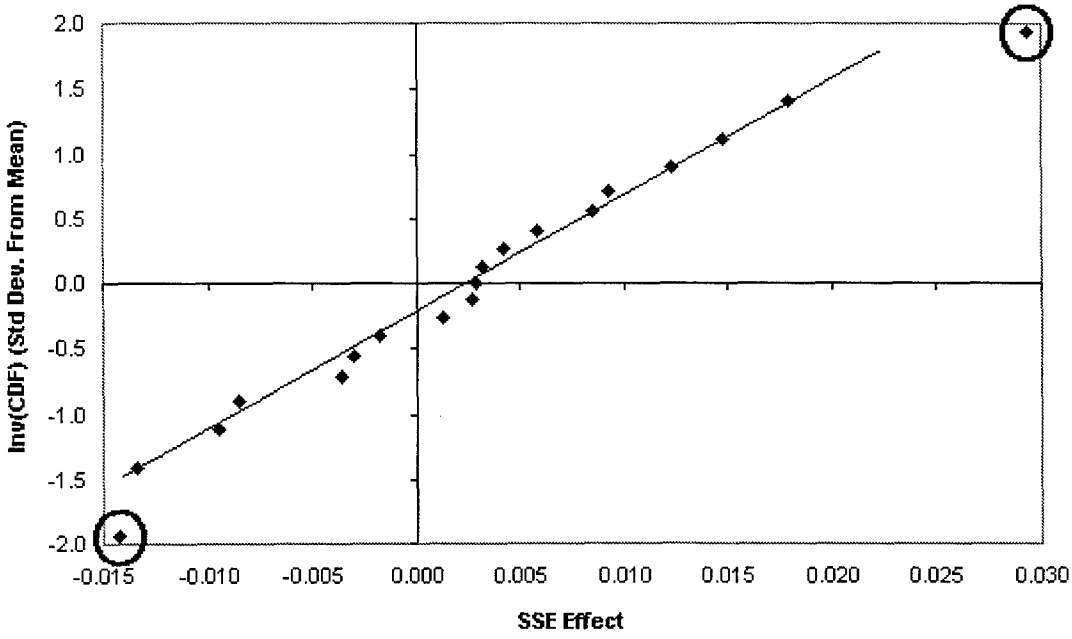


Figure 12. Experiment #1 nodes selected as a result of the initial screening.

Additional Screenings

To pinpoint the source, a second more refined area of interest was developed from the results of the first screening, as shown in Figure 13.

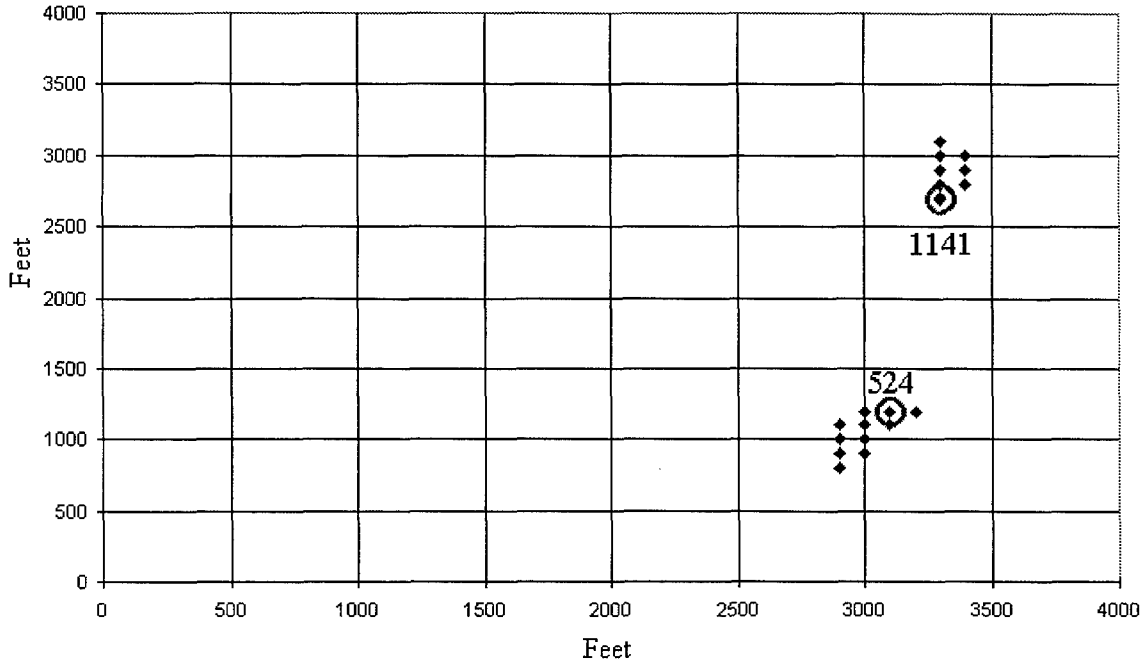


Figure 13. Experiment #1 locations used in second screening.

The 19 locations selected for the second screening were selected from those locations near (and including) those wells identified as significant in the first screening process. Figure 14 graphically identifies the most statistically significant wells selected by the second screening.

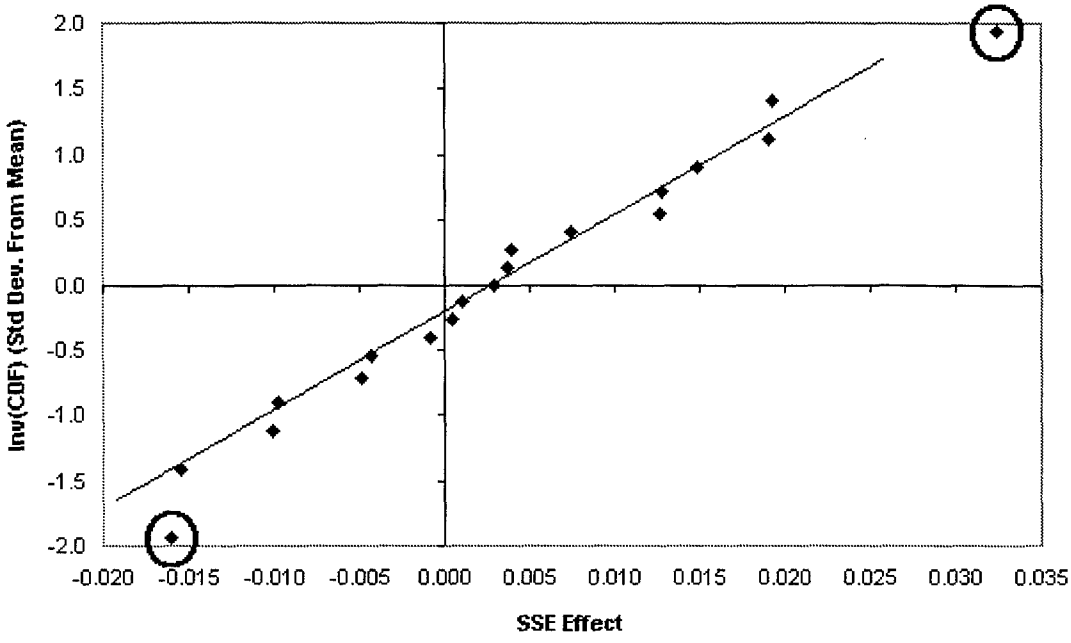


Figure 14. Experiment #1 results of second screening.

With the second screening complete, no further screening was needed since the locations used in the second screening were adjacent to each other in the numerical model. If finer resolution is required, the numerical model would be modified to include additional nodes between the existing nodes in the model. These additional nodes could then be used in a third screening to further refine the location of the contaminant source.

Part II -- Obtaining the Flux of the Located Sources

Once the statistically significant sources were found, the flux for each source was calculated using part two of the RSM process. This part required the straight-scaled high, low, and midpoint values for the flux ranges selected earlier, together with the *measured concentration values* at the original measurement wells.

Each iteration of the RSM process ideally improves the calculated flux at the screened source locations since this calculated flux value approaches the value needed to produce the *measured concentration values* of the initial measurement wells. In this experiment, five iterations of the RSM process were needed to reduce the SSE to within six significant digits.

Part two of the RSM process included a first stage that generates the reduction gradient and a second stage that employed multiples of the reduction gradient until the SSE increases on the response surface. Figures 15 shows graphically how the SSE value was reduced through several model runs and then increased in later runs.

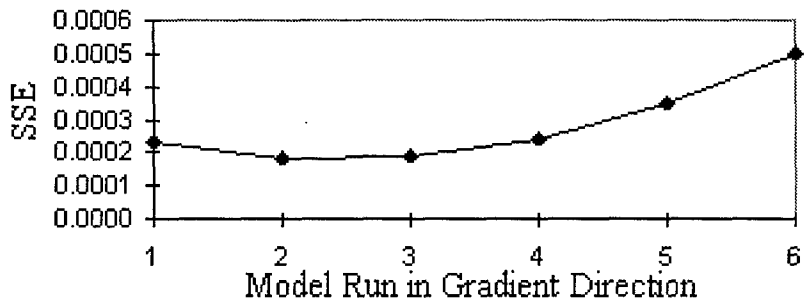


Figure 15. Experiment #1 divergence from SSE surface -- first iteration.

As the number of phase two model runs increased, the SSE values began to converge to a local minima, and the number of significant digits of accuracy increased. Figure 16 shows the results of the fifth and final phase II run, and identifies SSE values which are to six significant digits.

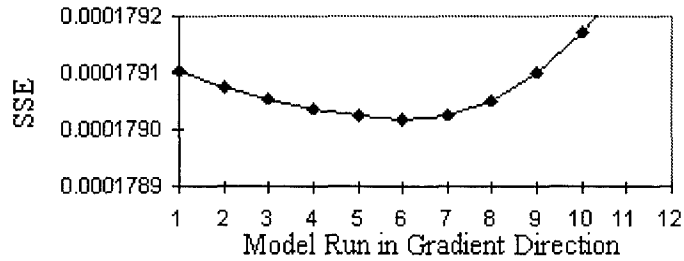


Figure 16. Experiment #1 divergence from SSE surface -- fifth iteration.

Completion of Process

Once the calculated SSE was within a predetermined number of significant digits of accuracy, the RSM process was complete. The flux obtained for the screened source nodes was obtained from the programmed flow rate values and is at the midpoint between the two limiting values calculated for each well used in phase 2.

Experiment Two

Introduction

Experiment two was designed to evaluate the effectiveness of the process given that a number of measurement wells currently exist on the site. In this experiment, the fuzzy processes developed earlier were not used and wells already in place at the Hill AFB OU3 site were modeled. This experiment illustrates the flexibility of RSM when measurement wells are selected for us.

The locations of the measurement wells were mapped to the numerical model and the SSE response surface was calculated from the difference between model run concentration values for these wells and the *measured values* of the alpha model. The modeled wells are shown in Figure 17.

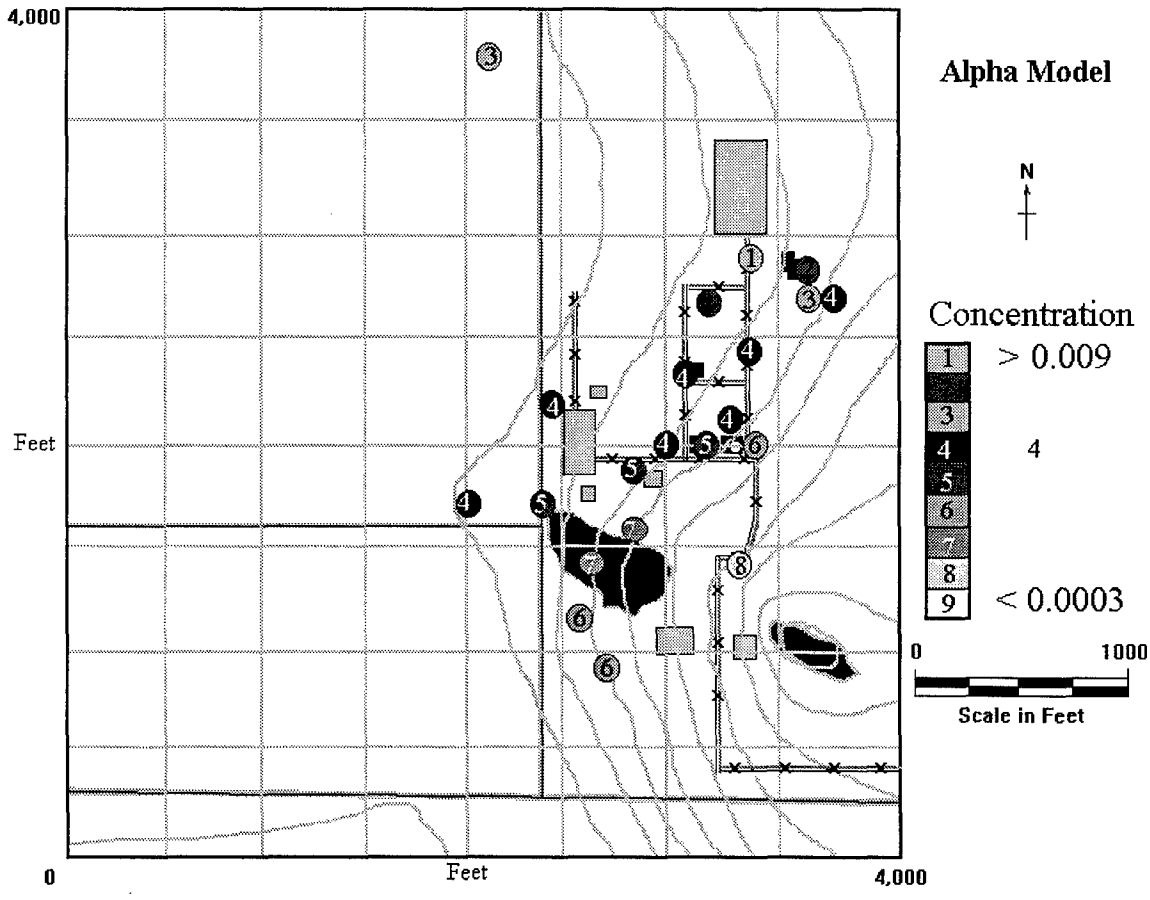


Figure 17. Experiment #2 well locations and concentrations.

Selecting Candidate Regions for the Migrated Contaminant

The five sites selected in experiment one were again selected as candidates from which a continuous contaminant source developed and migrated. However, instead of developing fuzzy *regions of possibility* around these sites (as in experiment number one), *areas of interest* were immediately developed by noting the location of these five sites and utilizing the levels of contaminant concentrations measured at the pre-existing wells. These areas of interest are shown in Figure 18.

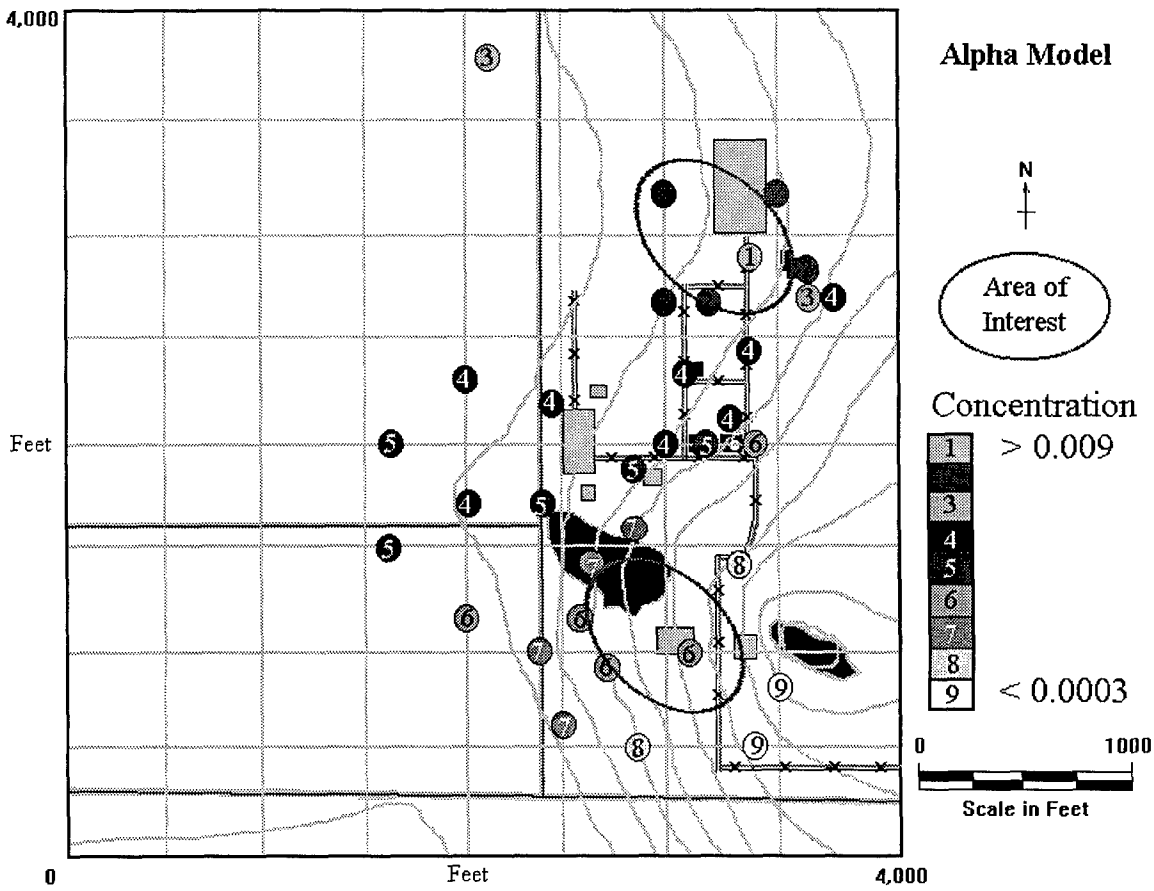


Figure 18. Experiment #2 bounding wells and areas of interest.

Screening for Significant Locations

The large number of locations represented by the selected areas of interest (ninety-nine locations, as shown in Figure 19) again prompted a repeated reduction from a large single run to two smaller runs.

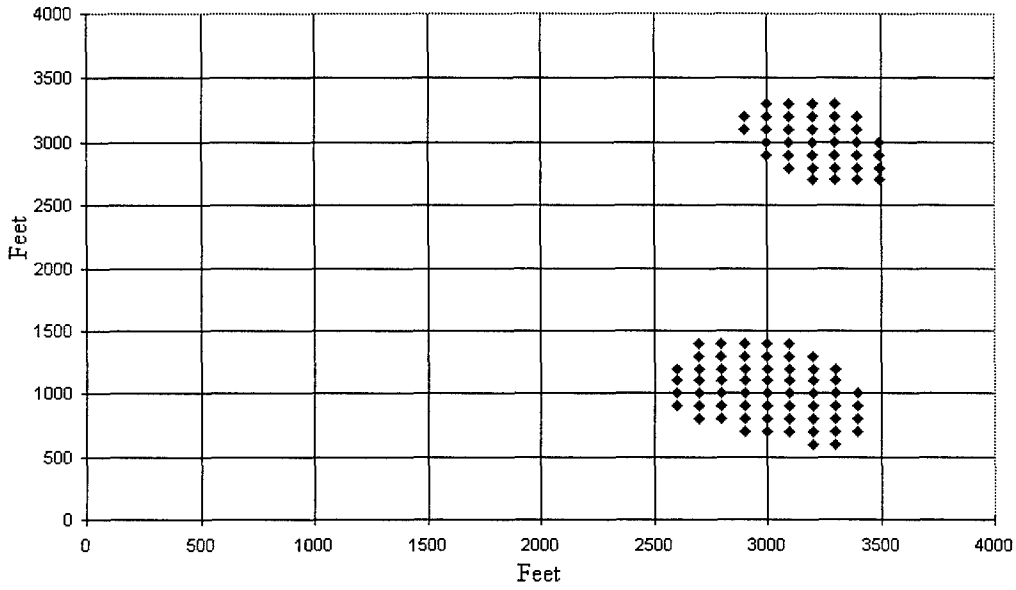


Figure 19. Experiment #2 candidate nodes for screening process.

Using a PB design for 20 runs (19 locations) solutions for the initial screening of experiment two were obtained as shown in Figures 20 and 21.

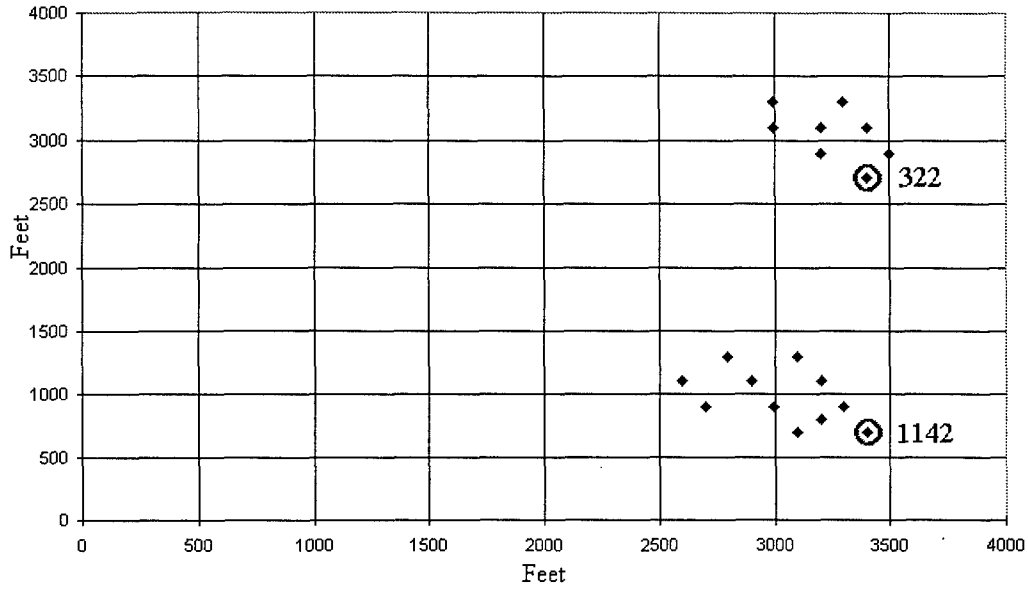


Figure 20. Experiment #2 nodes used and selected by initial screening.

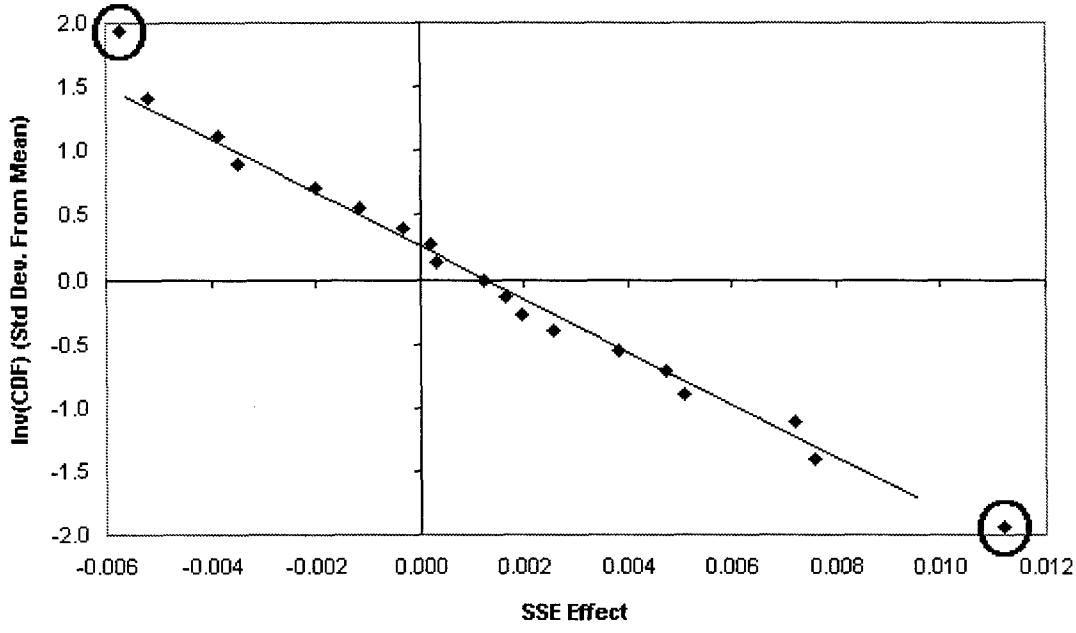


Figure 21. Experiment #2 results of initial screening.

Second Screening

A refined *area of interest* was developed from the results of the first screening. However, because of the proximity of the significant solutions, a reduced PB design that utilized only 11 locations for the second screening was used. These locations are shown in Figure 22, while Figure 23 graphically identifies the most statistically significant wells selected by the second screening in experiment two.

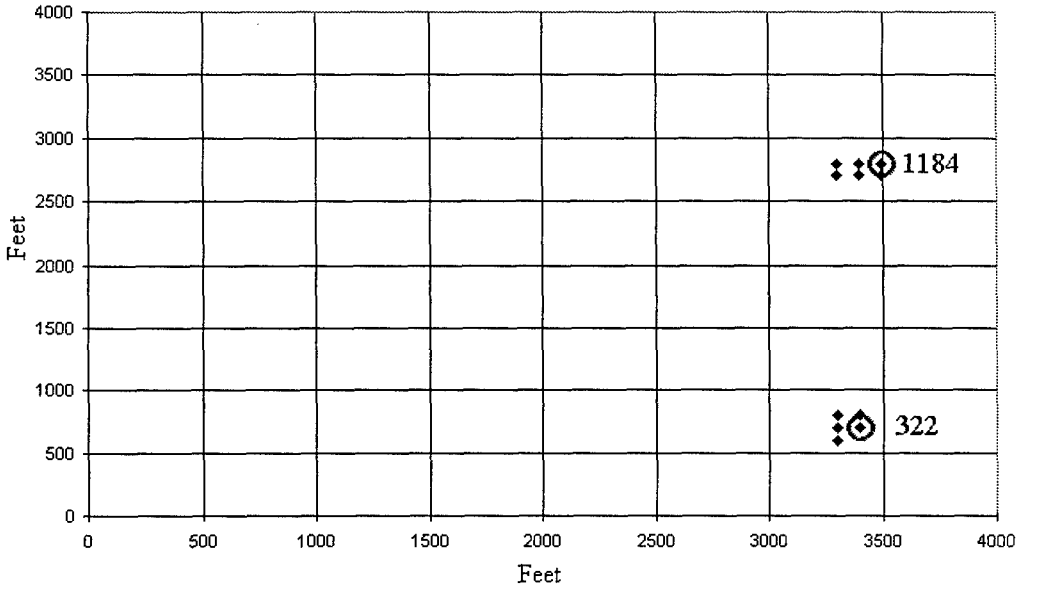


Figure 22. Experiment #2 locations used in second screening.

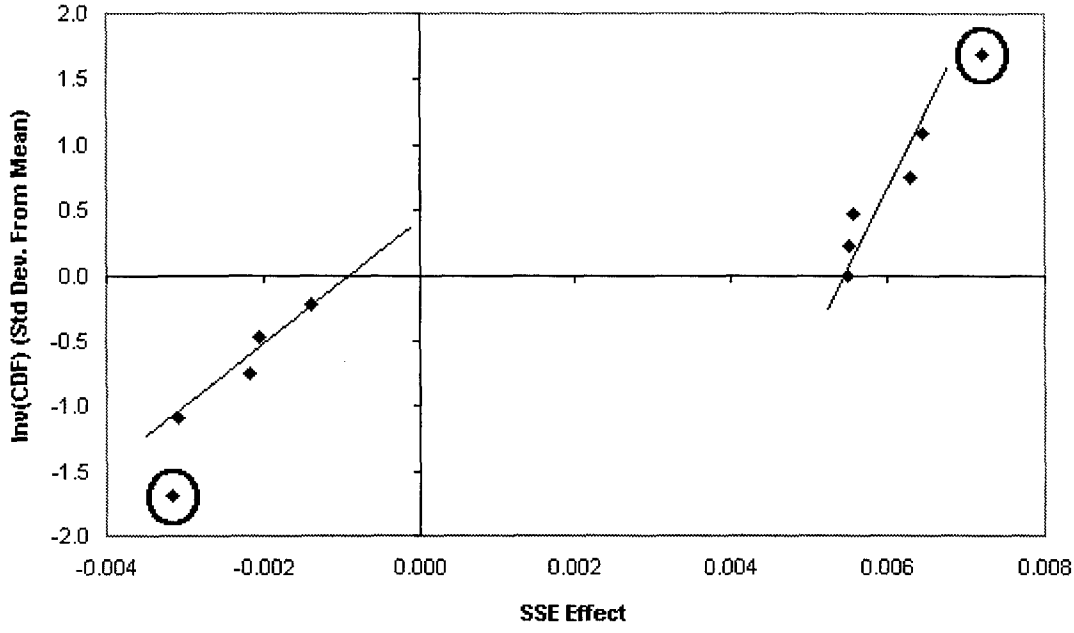


Figure 23. Experiment #2 results of second screening.

Part II -- Obtaining the Flow Rate of the Located Source(s)

Once the statistically significant sources were found, the flow rate for each source was calculated using part two of the Response Surface Methodology (RSM) process. Four iterations of the RSM process were needed to locate the (local) minimum value on the SSE response surface. Figure 24 shows graphically how the SSE value first decreases through several model runs and begins to increase in later runs.

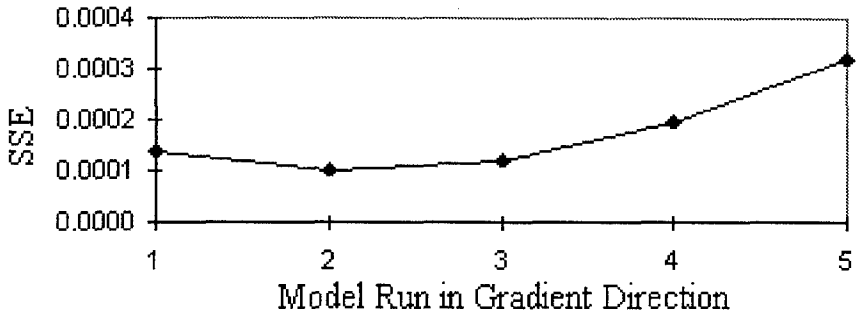


Figure 24. Experiment #2 divergence from SSE surface -- first iteration.

In this experiment, because the calculated flow rate value for one of the input parameters converged to an exact solution, the calculated reduction gradient for the fourth iteration consists of only a single-changing parameter, and the fourth set of runs focused solely on changing the single parameter's settings in the direction of steepest descent. Table 1 outlines the flow rate settings used in the fourth iteration of the RSM Part two process. Note that instead of successive multiples of the single-parameter reduction gradient, the reduction gradient difference was multiplied by an exponential value until a divergence from the response surface was observed. Figure 25 outlines the graphed results of this set of model runs.

TABLE 1.
EXP 2 RESULTS FROM FOURTH ITERATION OF RSM PROCESS.

Run Number	Low Zone Flow Rate Value	High Zone Flow Rate Value (Fixed)	SSE
1	0.0000496749	0.0003500450	0.000104474580
2	0.0000496250	0.0003500450	0.000104469583
3	0.0000495751	0.0003500450	0.000104464592
4	0.0000494749	0.0003500450	0.000104454588
5	0.0000492749	0.0003500450	0.000104434689
6	0.0000488749	0.0003500450	0.000104395172
7	0.0000480749	0.0003500450	0.000104317264
8	0.0000464749	0.0003500450	0.000104165948
9	0.0000432749	0.0003500450	0.000103881319
10	0.0000368749	0.0003500450	0.000103384075
11	0.0000240749	0.0003500450	0.000102677680
12	0.0000200000	0.0003500450	0.000102533416
13	0.0000150000	0.0003500450	0.000102409602
14	0.0000100000	0.0003500450	0.000102344419
15	0.0000050000	0.0003500450	0.000102337871
16	0.0000040000	0.0003500450	0.000102343598
17	0.0000030000	0.0003500450	0.000102351670
18	0.0000025000	0.0003500450	0.000102356586

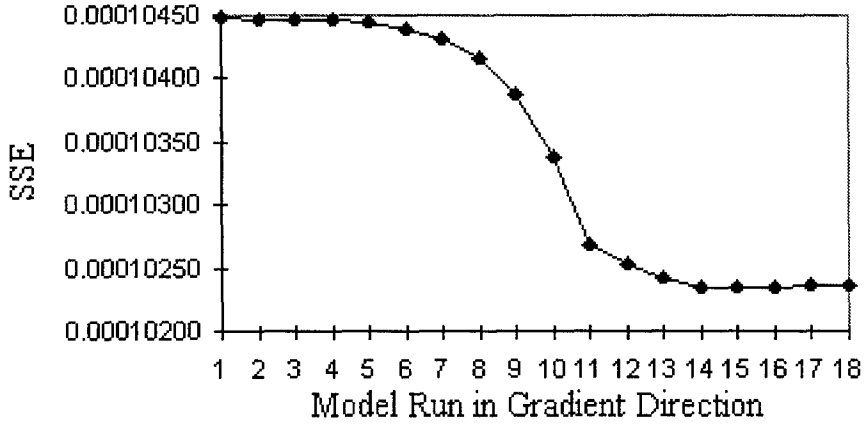


Figure 25. Experiment #2 divergence from SSE surface -- fourth iteration.

Completion of the RSM Process

As in the first experiment, once the calculated SSE was within a predetermined number of significant digits of accuracy, the RSM process was complete. The flux obtained for the screened source nodes was again obtained from the programmed flow rate values which were midpoint between the two limiting values calculated for each well used in phase 2.

Experiment Three

Introduction

Experiment three was designed to evaluate the effectiveness of the process given a scenario similar to experiment two, where a number of measurement wells currently exist on the site. However, the truth run was modified so model permeability values at each node are from a log normal distribution rather than from the same averaged value at all nodes. This new (beta) model tests whether averaging hydraulic conductivity values over a large area affects the results of the RSM process for a multiple contaminant source. This estimation is typical (and necessary) in practice when producing a numerical model of a realistic site like the alpha site. While the first two experiments tested the RSM procedure in a "data rich" environment, this experiment was designed to see if RSM would perform effectively when the model used was an approximation to the truth set.

A comparison of the distribution of hydraulic conductivity values used in experiment 3 is presented in Table 2 as compared with measurements obtained from the Borden (Sudicky, 1986), Cape Cod (Hess et al., 1992), and Columbus (Rehfeldt et al.,

1992) sites. Note that because of the log transformed nature of the data, the variance (and not the coefficient of variance) was used as the comparison value. The Beta Model is very representative of those actual sites, having a variance value of 0.67 cm/s, which is slightly larger than that found at the Borden and Cape Cod sites, and smaller than that of the Columbus site. The distribution of hydraulic conductivity values in the truth run makes approximating the set a realistic test case, yet more toward a "best case" (homogeneous) rather than a "worst case" (heterogeneous) scenario for the experiment.

TABLE 2.
EXP 3 BETA MODEL COMPARED WITH OTHER KNOWN SITES
(Adapted from Heiderscheidt, 1996)

Site Name	Mean of ln(K) (cm / s)	Variance of ln(K)
Beta Model	-11.55	0.67
Borden	-4.934	0.29
Cape Cod	-3.352	0.14
Columbus	-9.7	5.5

The locations of the measurement wells were again mapped to the numerical model and the SSE response surface was calculated from the difference between model run concentration values for these wells and the *measured values* of the beta model. The modeled wells are shown in Figure 26.

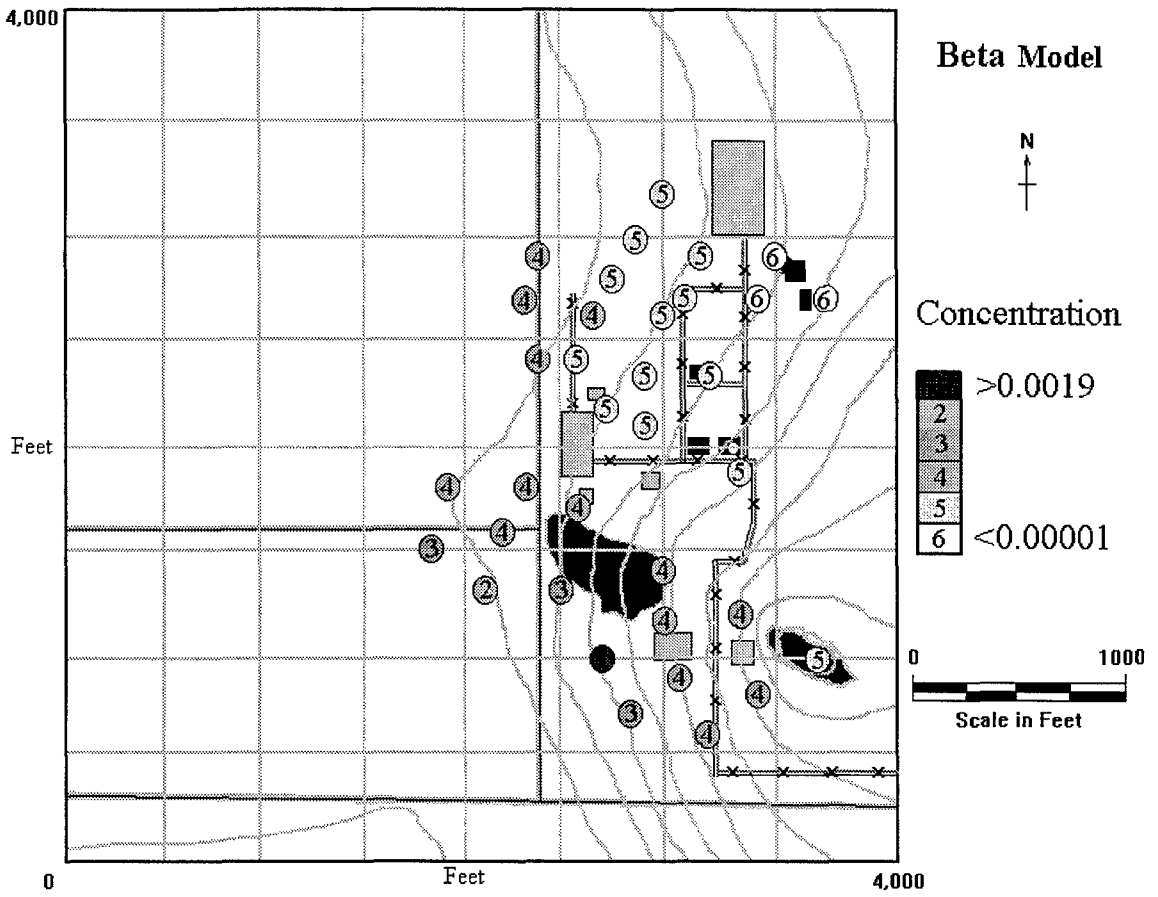


Figure 26. Experiment #3 well locations and relative concentrations.

Selecting Candidate Regions for the Migrated Contaminant

The five sites selected in experiments one and two as potential historical leak sources were again selected as candidates from which a continuous contaminant source developed and migrated. These regions of possibility and concentration values based on these potential leak sites are shown in Figure 27.

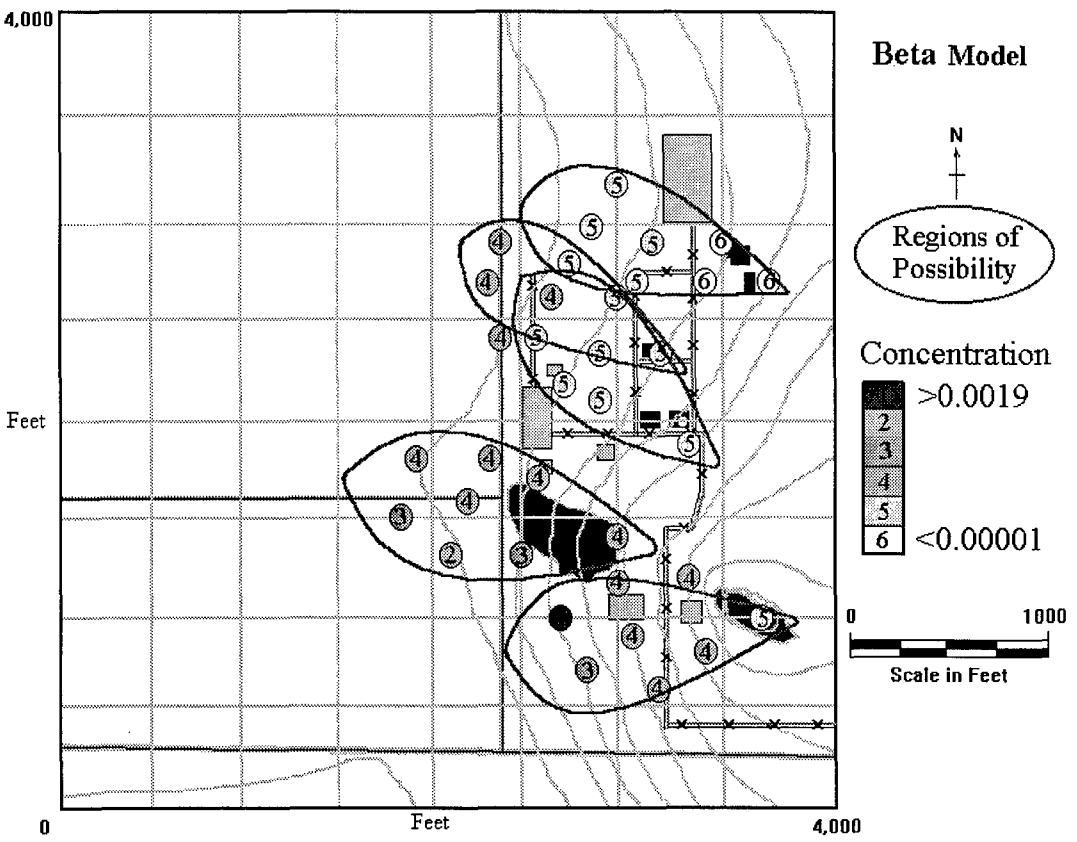


Figure 27. Experiment #3 regions of possibility developed from historical leak sites.

The areas of interest selected for these regions of possibility, developed from initial and additional bounding wells, are shown in Figure 28. Note that in this experiment, I again returned to the unplaced measurement well scenario, and used fuzzy logic to map a pattern of measurement wells.

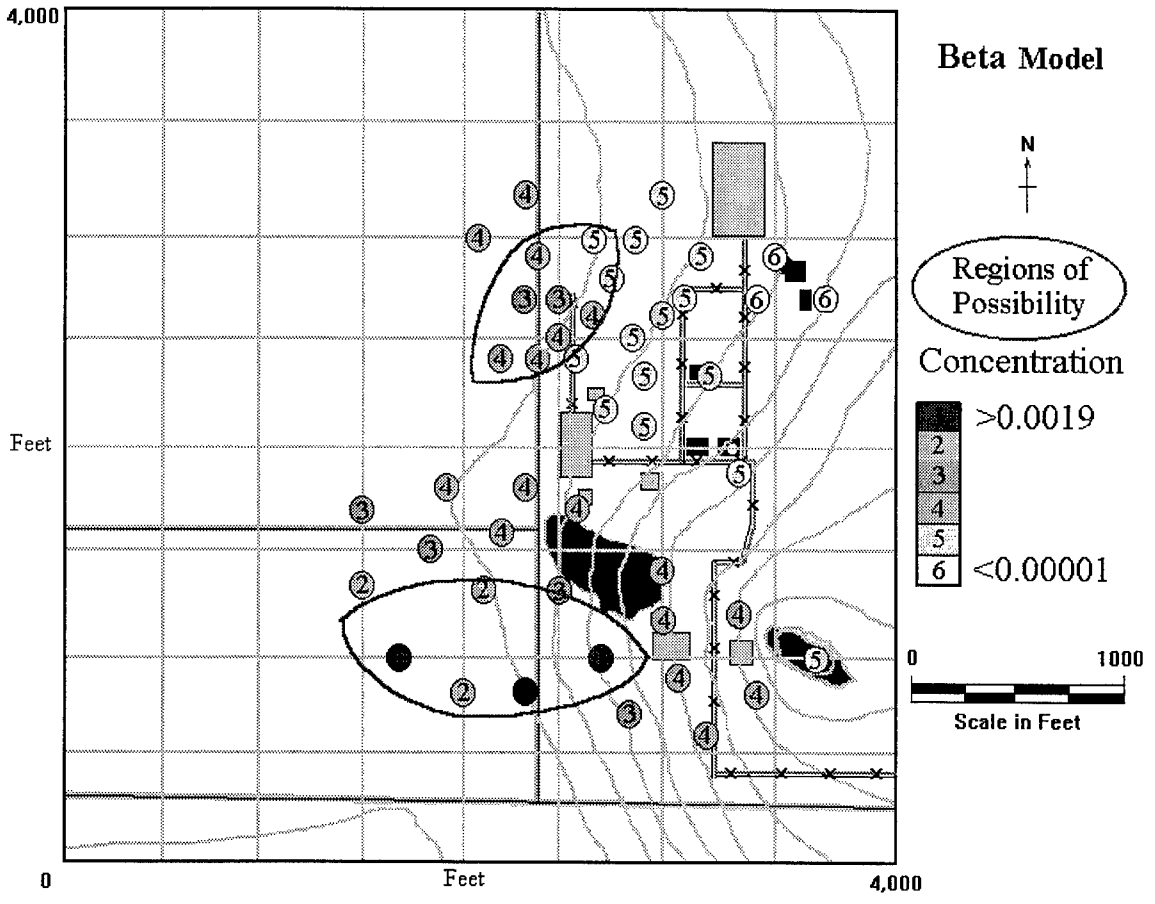


Figure 28. Experiment #3 bounding wells and areas of interest.

Screening for Significant Locations

The large number of locations represented by the selected areas of interest again prompted a repeated reduction from a large single run to two smaller runs.

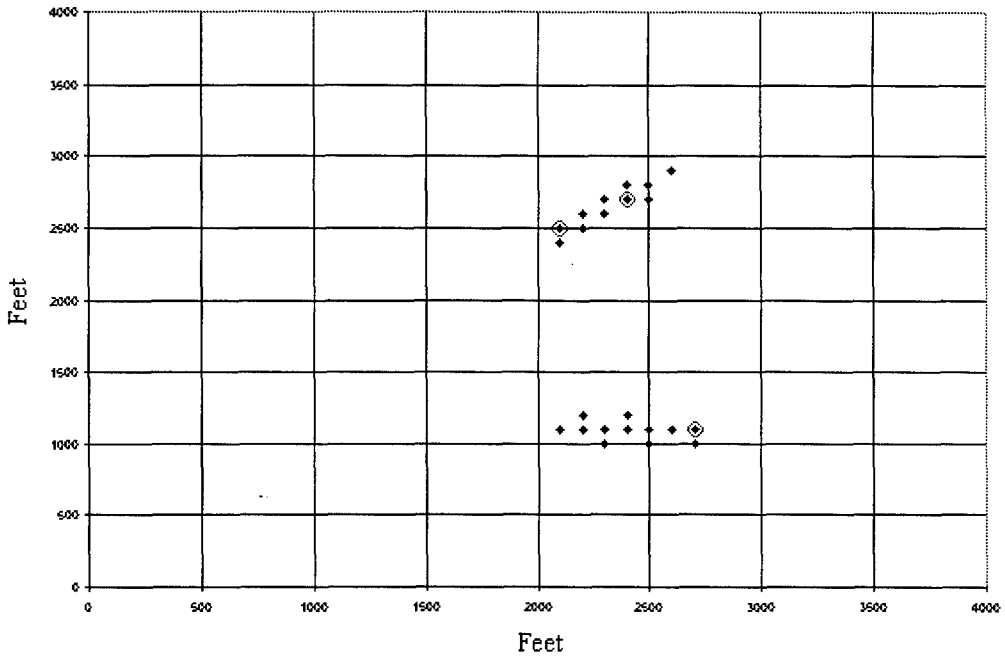


Figure 29. Experiment #3 locations used and significant nodes obtained.

Using PB designs similar to those in experiments one and two, the solution shown in Figure 29 was obtained for the final screening of experiment three. Once these sources were screened, the flow level for each identified source was again calculated using part two of the RSM process.

V. Results of Experiments

The RSM technique produced the results in Figure 30 and Table 3 for experiment one, Figure 31 and Table 4 for experiment two, and Figure 32 and Table 5 for experiment three.

In experiment one, the location of the first calculated source was (3300, 2700) on the finite element grid. For this single point source contaminant, the RSM process calculated a source approximately 300 feet (3 nodes) from the actual source at (3400, 3000) as shown in Figure 30. The second calculated source was located at (3100, 1200), 100 feet (1 node) from (3100, 1100), the nearest adjacent node of nine in the 3x3 contaminant source.

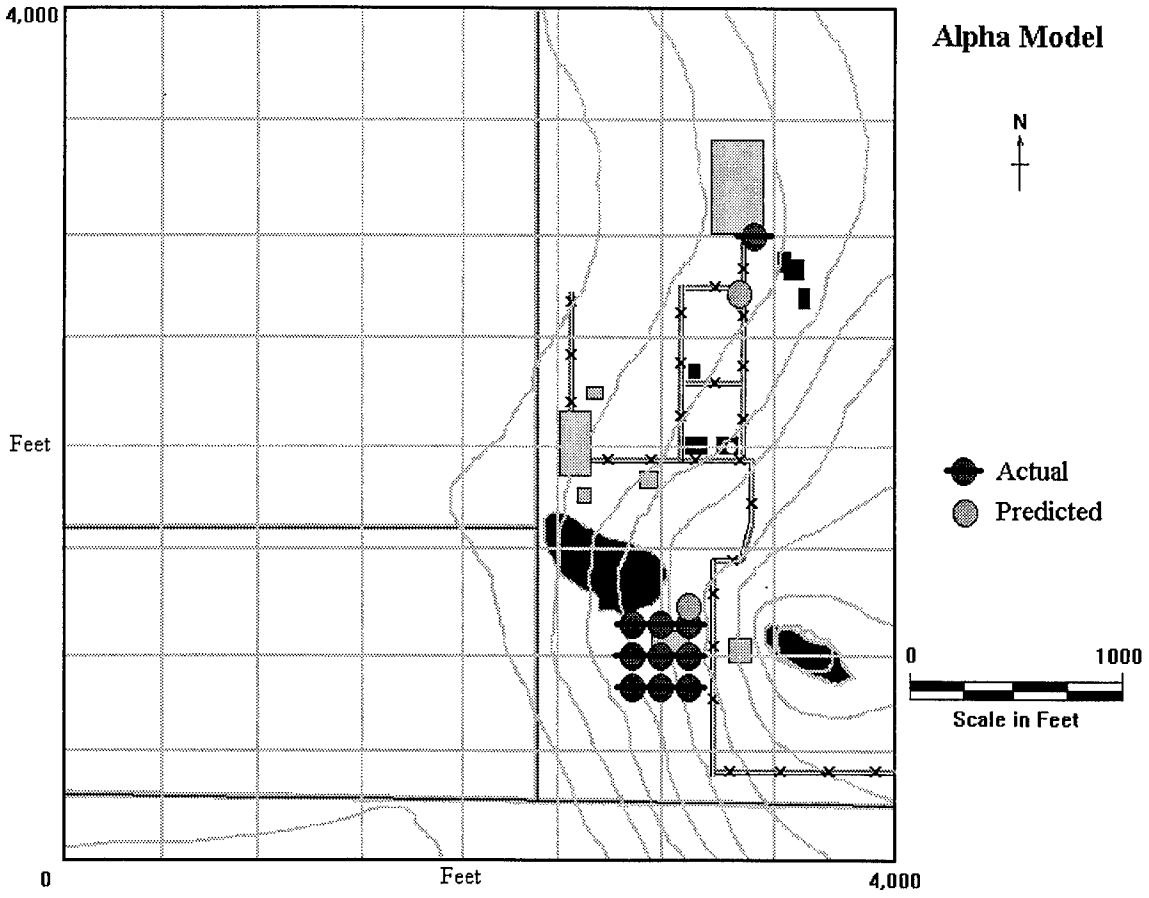


Figure 30. Experiment #1 true vs. calculated location of contaminant.

TABLE 3
RESULTS OF EXPERIMENT ONE

Actual Node Number	Actual Location (ft)	Actual Flow Rate (m^3/day)	Calculated Location	Calculated Flow Rate m^3/day
399	2900, 900	1.7×10^{-6}	NA	NA
400	3000, 900	1.7×10^{-6}	NA	NA
401	3100, 900	1.7×10^{-6}	NA	NA
440	2900, 1000	1.7×10^{-6}	NA	NA
441	3000, 1000	1.7×10^{-6}	NA	NA
442	3100, 1000	1.7×10^{-6}	NA	NA
481	2900, 1100	1.7×10^{-6}	NA	NA
482	3000, 1100	1.7×10^{-6}	NA	NA
483	3100, 1100	1.7×10^{-6}	3100, 1200	5.0×10^{-5}
1265	3400, 3000	4.4×10^{-4}	3300, 2700	3.6×10^{-4}

In experiment two, the location of the first calculated source was (3500, 2800) on the finite element grid as shown in Figure 31.

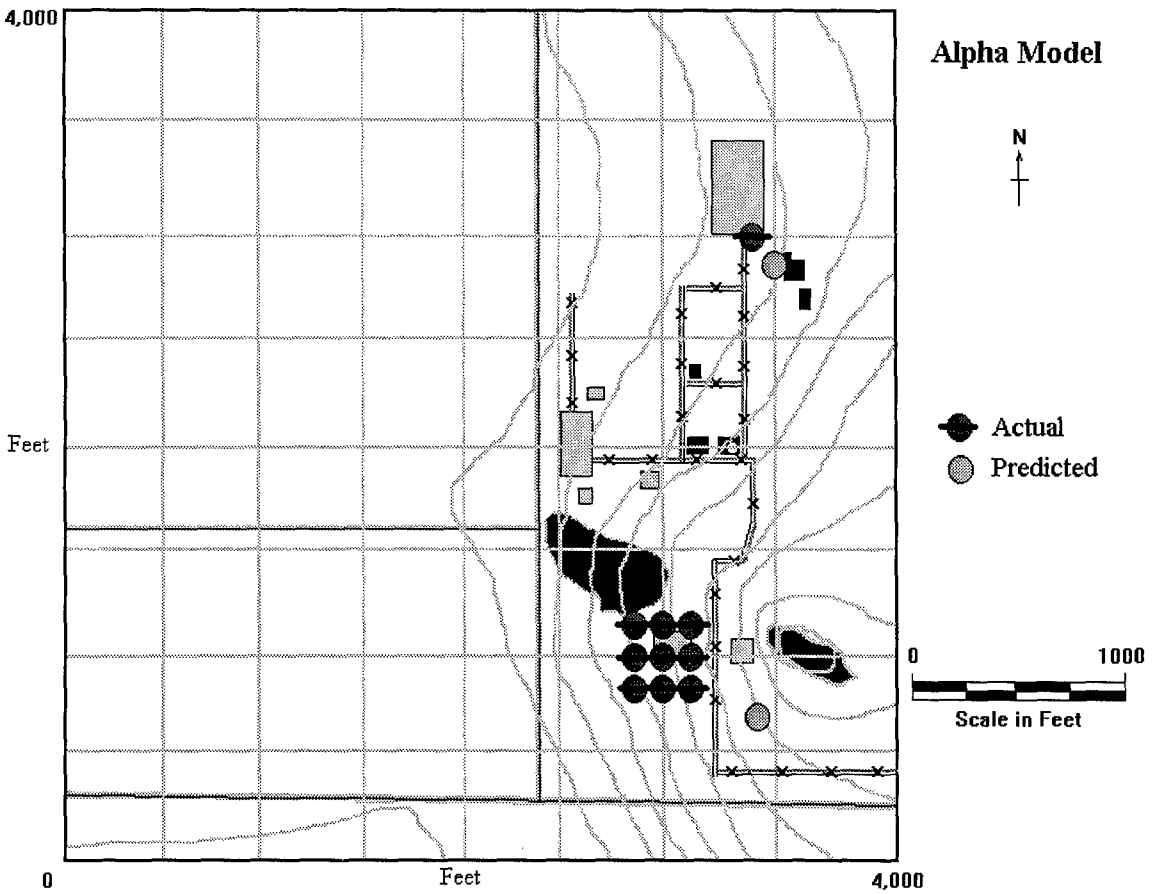


Figure 31. Experiment #2 true vs. calculated location of contaminant.

For this single point source contaminant, the RSM process calculated a contaminant source at (3500, 2800), approximately 225 feet (2.25 nodes) upstream from the actual source at (3400, 3000). The second calculated source was located at (3400, 700), approximately 360 feet upstream (3.6 nodes) from the nearest node of the actual source at (3100, 900).

TABLE 4
RESULTS OF EXPERIMENT TWO

Actual Node Number	Actual Location (ft)	Actual Flow Rate m^3/day	Calculated Location	Calculated Flow Rate m^3/day
399	2900, 900	1.7×10^{-6}	NA	NA
400	3000, 900	1.7×10^{-6}	NA	NA
401	3100, 900	1.7×10^{-6}	NA	NA
440	2900, 1000	1.7×10^{-6}	NA	NA
441	3000, 1000	1.7×10^{-6}	NA	NA
442	3100, 1000	1.7×10^{-6}	NA	NA
481	2900, 1100	1.7×10^{-6}	NA	NA
482	3000, 1100	1.7×10^{-6}	NA	NA
483	3100, 1100	1.7×10^{-6}	3400, 700	5.0×10^{-6}
1265	3400, 3000	4.4×10^{-4}	3500, 2800	3.5×10^{-4}

In experiment three, the location of the first calculated source was (2700,1100) on the finite element grid as shown in Figure 32, located approximately 225 feet (2.25 nodes) upstream from the actual source at (2500,1000). The second calculated source was located at (2100,2500), approximately 500 feet (5 nodes) downstream from (2600, 2600), the nearest node of four in the leaky pipe source.

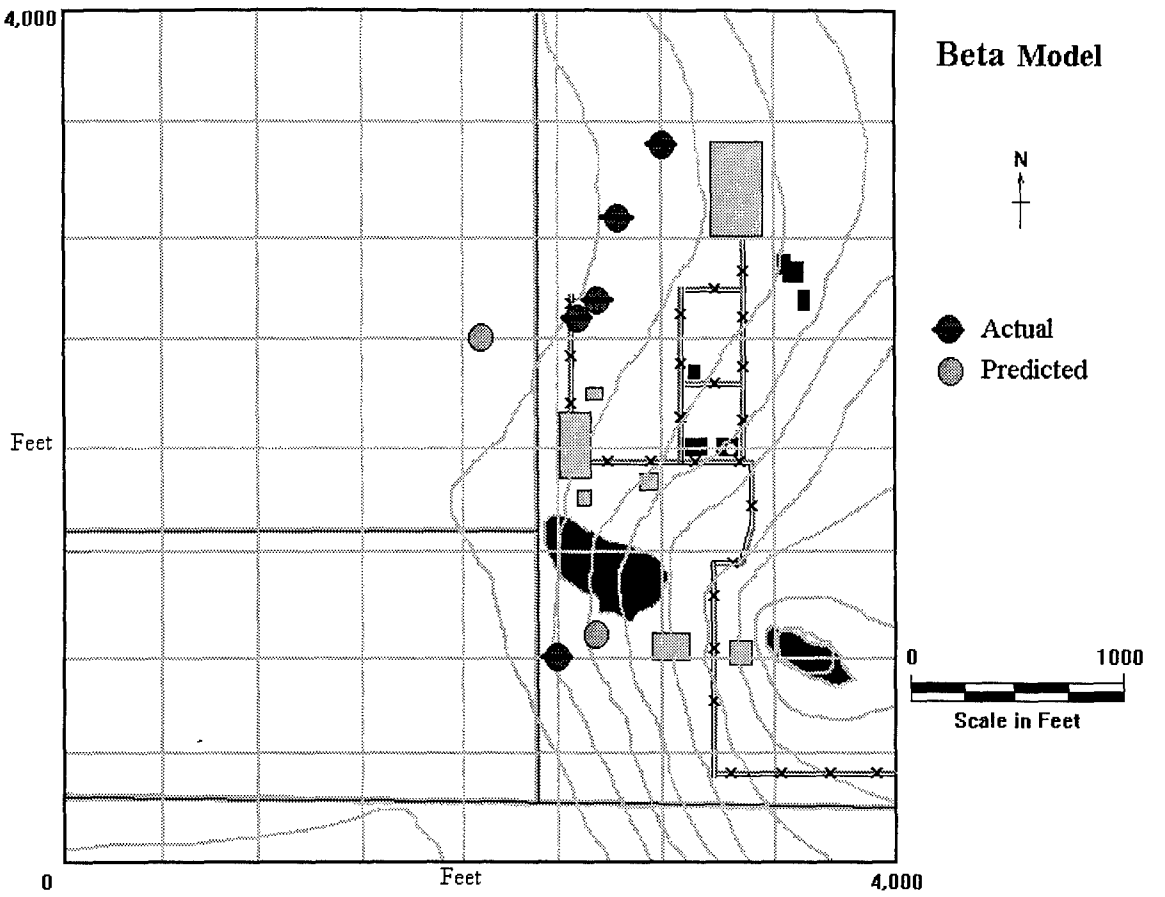


Figure 32. Experiment #3 true vs. calculated location of contaminant.

TABLE 5
RESULTS OF EXPERIMENT THREE

Actual Node Number	Actual Location (ft)	Actual Flow Rate m^3/day	Calculated Location	Calculated Flow Rate m^3/day
436	2500, 1000	1.5×10^{-4}	2700, 1100	1.2×10^{-4}
1092	2600, 2600	1.2×10^{-6}	2100, 2500	3.4×10^{-5}
1135	2700, 2700	1.2×10^{-6}	NA	NA
1300	2800, 3100	1.2×10^{-6}	NA	NA
1425	3000, 3400	1.2×10^{-6}	NA	NA

Part of the reason for the identification of the particular source node for the lower-concentration area in experiment three is the inexactness of the RSM method. It appears that when parameters in adjacent nodes are so close together in magnitude that the

differences between the SSEs produced by them are small, the screening plot is difficult to analyze and the potential for selecting the incorrect source node increases. Figure 33 shows an example of this phenomena.

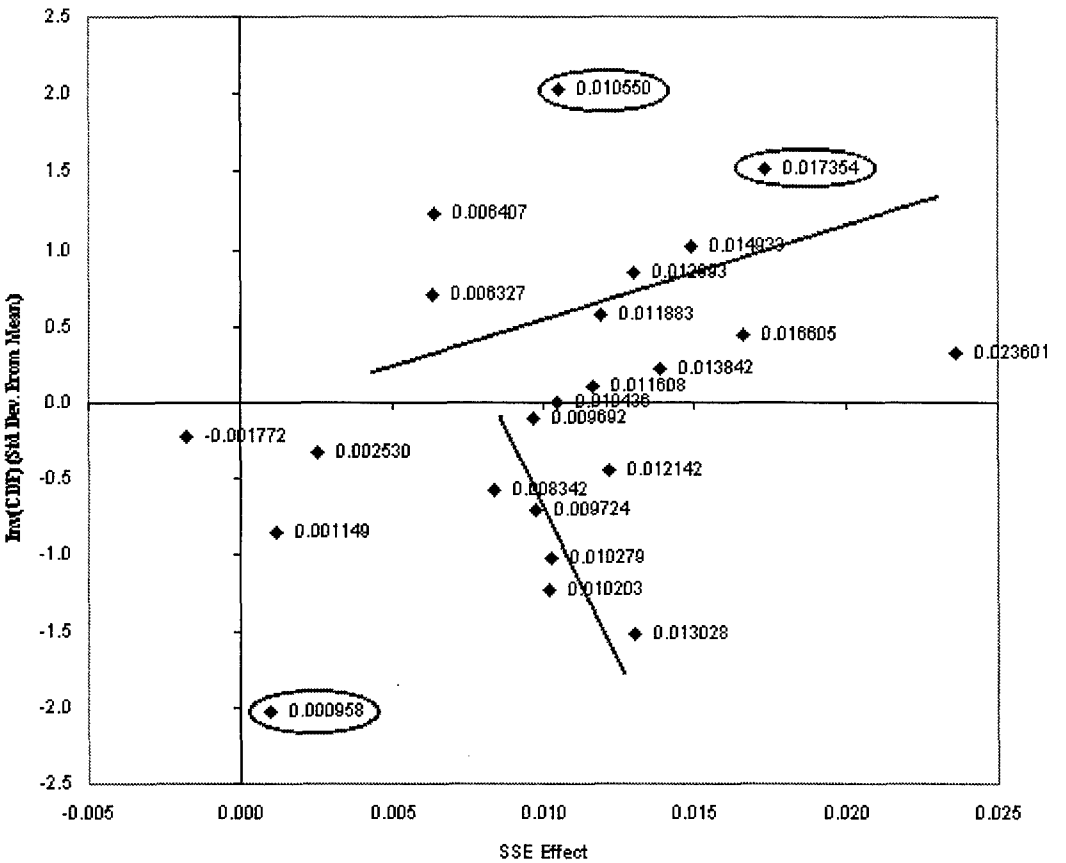


Figure 33. Inexactness of screening contributes to incorrect selection of node.

Summary of Results

In experiments one and two, although the location and contaminant flow rate values obtained for the upper source were within 225 - 300 feet (2 to 3 nodes), the reason for the lower source flux values at these locations can be attributed to the relative location

of the calculated source with respect to the majority of the measurement wells. Since the calculated source was closer than that of the actual source, a lower source flux was needed at that closer source to minimize the value of the SSE. Likewise, the concentration source being closer to the lower concentration zone required a smaller calculated source flux for the lower zone to minimize the value of the SSE. The resulting source flux for the lower concentration source is also very close, if one considers the 9:1 ratio of actual to calculated sources. This consideration produces a result for the calculated lower-concentration source that is within an order of magnitude of the actual nine-location spread source for both experiments.

In experiment three, the location and contaminant mass flux values calculated for the higher-concentration source using RSM were very close (225 feet and same order of magnitude) to the truth values. However, the calculated lower-concentration source flux value and location had a larger error from the actual (500+ feet distant and 25% of actual flux). This error can be attributed to the relative location of the calculated source with respect to the majority of the measurement wells. Since the calculated source was farther from the wells being used to calculate the error slope, a higher source flux was needed at that farther source to minimize the value of the SSE. We found that during the RSM phase of the third experiment, the source flux value of one of the two sources identified as significant in the lower concentration zone caused a further minimization of the SSE by going to zero. This source was therefore eliminated as a possibility for the actual source of the leak. However, the source eliminated was closer to the actual source than the

remaining source in the lower concentration zone. Further investigation on how screening techniques are affected by nodes which have a relatively small difference in error is an area recommended for future study.

VI. Conclusions and Recommendations

Conclusions

RSM is a useful tool for gaining fundamental knowledge about a contaminated site. Through the use of existing measurement data and other information typically obtained during early stages of the site characterization process, it can be used to provide hydrologists with a reasonable initial estimate of a contaminant's location and mass flux when the contaminant is modeled as a point source. This systematic method has the potential for providing a quick, statistically accurate solution to the groundwater contaminant inverse problem using most groundwater and contaminant transport models.

RSM permits hydrologists to examine and modify multiple parameters simultaneously, allowing them to obtain valuable information about the contaminant while taking into account the interactive effects of different input parameters. Methods which employ single-parameter adjustments are likely to miss such effects. RSM can also be used when estimating flow and transport model parameters such as hydraulic conductivity, which is typically log normally distributed.

The results in this work were obtained using multiple sources and hydraulic conductivity variances representative of a homogeneous site, yet can be applied to sites with variances of a more heterogeneous nature.

The accuracy of RSM in practice is dependent upon the accuracy of the conceptual model, with the quality of the estimate of the contaminant's location and mass flux dependent on how closely the numerical model maps to the actual site.

Recommendations

Continued research into applying RSM techniques to site remediation technology should include the following:

1. Examining the use of RSM on differing types of migrating contaminants. Although this work focused on contaminants with fixed mass flux concentration values, future work could investigate the usefulness of RSM to estimate the location of migrated contaminants given contaminants with pulse and time-variant source characteristics.

2. Determining whether RSM can be used to locate either the existing location of a contaminant or the initial historical source of the contaminant given a lack of site historical information and contaminant levels. (In this work the existence of several locations where initial potential sources of contaminant leaks and spills was assumed.)

3. Determining the effectiveness of RSM given non steady-state or unknown flow conditions.

4. Determine the effectiveness of RSM when considering more than just advective and dispersive transport, e.g., biodegration, sorption, etc.

5. Research recent theoretical advances in methods for locating source contaminants (such as fractal models, MODFLOWP, etc.) and compare RSM theory and results with this research.

6. Determine whether advances made in other disciplines, such as genetic algorithms and pattern recognition subfields within Artificial Intelligence, provide results comparable to RSM.

Appendix A. Calculating Values for SUTRA Model

To calculate the values needed for the pressure difference and permeability in the experiment, Darcy's Law was used to obtain the relationship between flux (flow rate) and head gradient, as shown in the equation (note continuation of equation numbers):

$$q = -K * dh / dx, \tag{7}$$

where K = Hydraulic conductivity in m/sec

q = Flux in m/sec from $Lw^3 / ((Lm^2) * t)$

dh / dx = Hydraulic head gradient

Lw = Length of water in meters

Lm = Length of porous media in meters

Velocity of water was then obtained by equation:

$$V = q / \theta_W, \tag{8}$$

where V = Velocity = distance along medium/time = Lm / T in m / sec

$\theta_W = Lw^3 / Lm^3$

Substituting (7) into (8) yields:

$$V = -(K / \theta_W) * dh / dx \tag{9}$$

Rewriting this equation and solving for K yields:

$$K = -V * \theta_W / (dh / dx) \tag{10}$$

Total hydraulic head is usually composed of two components. These two components are pressure head and elevation head. Although there are other components,

for this case they are not important. Because SUTRA incorporates unsaturated flow, and since it separates pressure head and elevation because θ_W is dependent on pressure head and not elevation,

$$h = P / (Y * g) + z, \quad (11)$$

where h = Hydraulic head

P = Pressure

z = Elevation (m)

Y = Density of water in kg / m^3

g = Gravitational acceleration in m / sec^2

Since dz / dx = 0 (elevation does not change with horizontal distance,) the average gradient was determined across the system by realizing the two constant pressure boundaries in the system. Therefore,

$$dh / dx\ (avg) = 1.0 / (Yg) * [\ P(x = L) - P(x = 0)] / (L - 0) \quad (12)$$

where L = Distance between measured elevation lines in meters

Rewriting (12) yields:

$$[P(x = L) - P(x = 0)] = (dh / dx) * L * Y * g \quad (13)$$

Which yields the Pressure Difference needed for the SUTRA Model

Now hydraulic conductivity is dependent on three factors; pore geometry, fluid properties, and saturation, and was estimated using:

$$K = k * Y * g / u, \quad (14)$$

where K is hydraulic conductivity

k = Permeability of water in meters

u = Kinematic viscosity in kg / (m sec)

Note that hydraulic conductivity in equation (14) is *saturated* conductivity, that is, all voids are filled with the fluid, and fluid properties are accounted for by u, Y, and g. The intrinsic permeability, k, is a property of the porous medium alone, which is important if dealing with the flow of several different fluids. (14) can then be rewritten as

$$k = K * u / (Y * g) \qquad (15)$$

Equations (13), (14) and (15) provided the values needed to complete the SUTRA model of the site. The values used in the SUTRA model are identified in tables A-1 and A-2.

TABLE A-1
MEASURED VALUES/CONSTANTS USED IN SUTRA MODEL

VARIABLE	VALUE
dh / dx	4750 - 4650 / 4000 = 0.025
V	1.0 m / day = 1.157×10^{-5} m / sec
θ_w	0.2
L	1219 m
Y	1000 kg / m^3
g	9.8 m / sec^2
u	10 kg / (m sec)

TABLE A-2
CALCULATED VALUES USED IN SUTRA MODEL

VARIABLE	Equation	VALUE
P	(13)	3.0E5 kg / (m sec^2)
K	(14)	9.256×10^{-5} m / sec
k	(15)	9.445×10^{-8}

Appendix B. Experimental Design Settings and Responses

This appendix contains the worksheets used to record the parameter settings and the associated responses for experiment one. The worksheets are arranged as follows:

* Note spacing of FORTRAN input files modified so data can be represented on page.

RSM Experiment One
Developing Regions of Possibility and Areas of Interest

Wells Generated by Fuzzy Logic

Well	X Loc.	Y Loc.	Node	Conc.
A-001	3700	2700	1145	0.001548
A-002	3400	2700	1142	0.003093
A-003	3500	2900	1225	0.006247
A-004	3100	2700	1139	0.004686
A-005	3200	2900	1222	0.009301
A-006	2900	3000	1260	0.007770
A-007	3000	3200	1343	0.006542
B-001	3200	2300	976	0.000993
B-002	2900	2300	973	0.001465
B-003	3000	2600	1097	0.003598
B-004	2600	2400	1011	0.002480
B-005	2700	2600	1094	0.004125
B-006	2800	2800	1177	0.006285
B-007	2300	2700	1131	0.003924
B-008	2400	2900	1214	0.005253
C-001	3300	1900	813	0.000220
C-002	2900	2100	891	0.000755
C-003	2700	2200	930	0.001277
C-004	2400	2400	1009	0.002560
D-001	3000	1400	605	0.000136
D-002	2500	1300	559	0.000251
D-003	2600	1700	724	0.000322
D-004	2100	1300	555	0.000289
D-005	2200	1600	679	0.000379
D-006	2300	1800	762	0.000543
D-007	1800	1500	634	0.000417
D-008	1900	1800	758	0.000676
E-001	3700	1000	448	0.000027
E-002	3400	800	363	0.000054
E-003	3300	1200	526	0.000095
E-004	3200	600	279	0.000046
E-005	3100	900	401	0.000189
E-006	3000	1200	523	0.000207
E-007	2800	700	316	0.000125
E-008	2700	1000	438	0.000304

The data on this page represents those nodes selected for the Regions of Possibility outlined in the theory section.

The wells are identified with a letter prefix for each region of possibility, starting from the top of the map and working down toward the bottom of the map. The number of the well is the identifier in that particular region.

The node number is the numerical model node number, which is derived from a 40x40 matrix (yielding effectively 1681 nodes.)

The 35 nodes represented here represent 2% of the total nodes in the numerical model.

RSM Experiment One
Developing Regions of Possibility and Areas of Interest

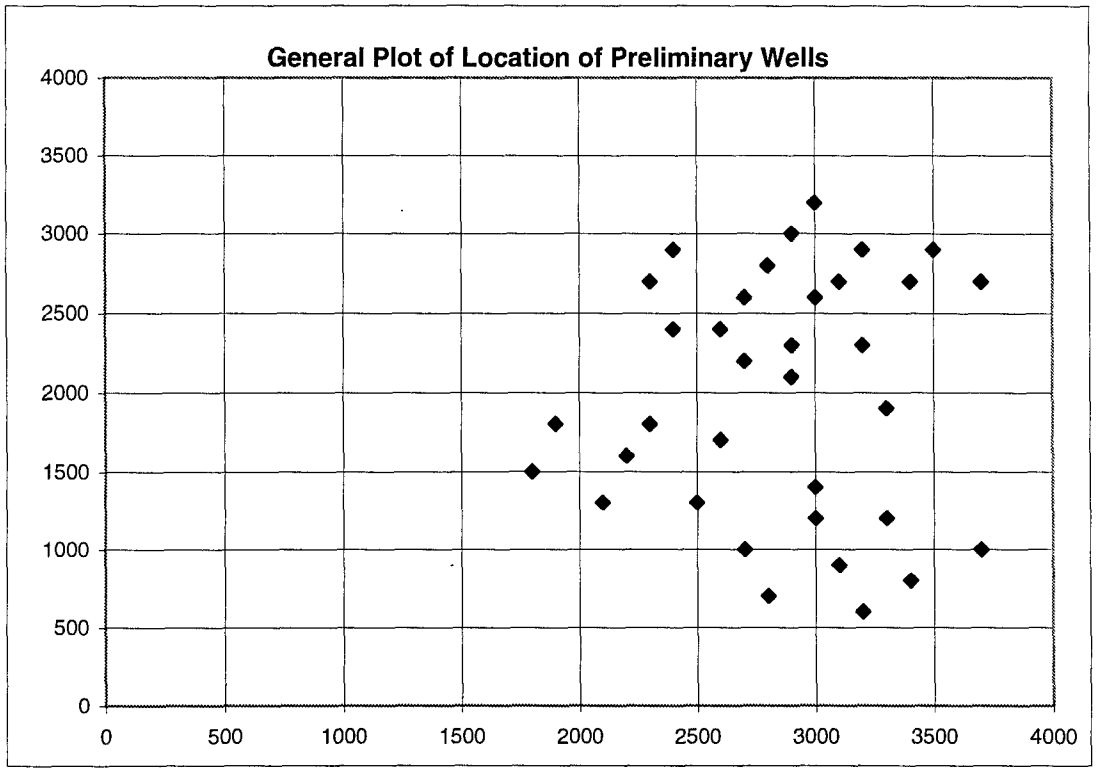
Preliminary Well Concentrations (Sorted by Concentration)

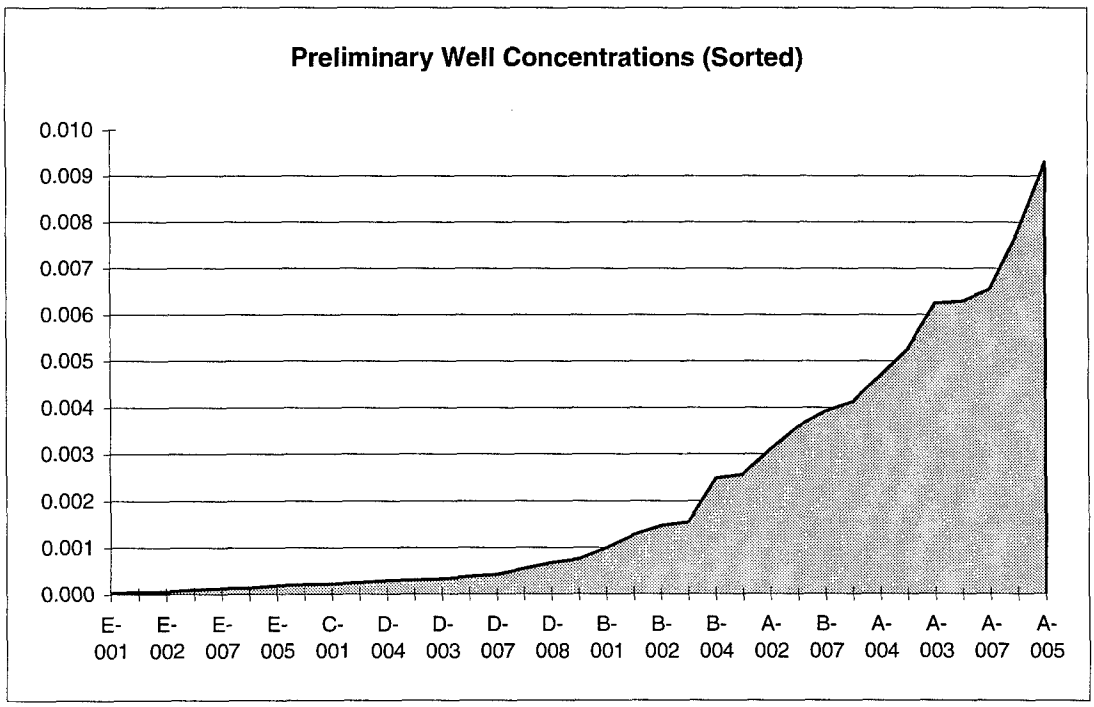
Well	X Loc.	Y Loc.	Node	Conc.
E-001	3700	1000	448	0.000027
E-004	3200	600	279	0.000046
E-002	3400	800	363	0.000054
E-003	3300	1200	526	0.000095
E-007	2800	700	316	0.000125
D-001	3000	1400	605	0.000136
E-005	3100	900	401	0.000189
E-006	3000	1200	523	0.000207
C-001	3300	1900	813	0.000220
D-002	2500	1300	559	0.000251
D-004	2100	1300	555	0.000289
E-008	2700	1000	438	0.000304
D-003	2600	1700	724	0.000322
D-005	2200	1600	679	0.000379
D-007	1800	1500	634	0.000417
D-006	2300	1800	762	0.000543
D-008	1900	1800	758	0.000676
C-002	2900	2100	891	0.000755
B-001	3200	2300	976	0.000993
C-003	2700	2200	930	0.001277
B-002	2900	2300	973	0.001465
A-001	3700	2700	1145	0.001548
B-004	2600	2400	1011	0.002480
C-004	2400	2400	1009	0.002560
A-002	3400	2700	1142	0.003093
B-003	3000	2600	1097	0.003598
B-007	2300	2700	1131	0.003924
B-005	2700	2600	1094	0.004125
A-004	3100	2700	1139	0.004686
B-008	2400	2900	1214	0.005253
A-003	3500	2900	1225	0.006247
B-006	2800	2800	1177	0.006285
A-007	3000	3200	1343	0.006542
A-006	2900	3000	1260	0.007770
A-005	3200	2900	1222	0.009301

Separation of Upper and Low Concentration Zones (Sorted by Concentration)

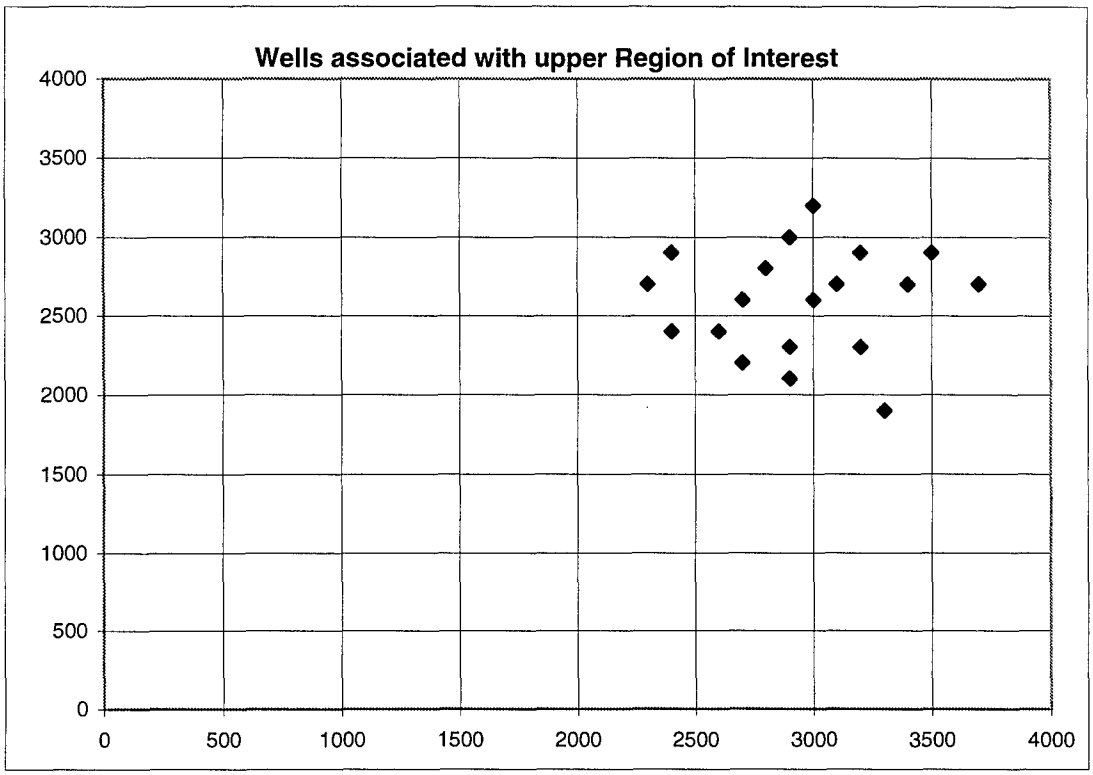
Well	X Loc.	Y Loc.	Node	Conc.	U or L
C-001	3300	1900	813	0.000220	U
C-002	2900	2100	891	0.000755	U
B-001	3200	2300	976	0.000993	U
C-003	2700	2200	930	0.001277	U
B-002	2900	2300	973	0.001465	U
A-001	3700	2700	1145	0.001548	U
B-004	2600	2400	1011	0.002480	U
C-004	2400	2400	1009	0.002560	U
A-002	3400	2700	1142	0.003093	U
B-003	3000	2600	1097	0.003598	U
B-007	2300	2700	1131	0.003924	U
B-005	2700	2600	1094	0.004125	U
A-004	3100	2700	1139	0.004686	U
B-008	2400	2900	1214	0.005253	U
A-003	3500	2900	1225	0.006247	U
B-006	2800	2800	1177	0.006285	U
A-007	3000	3200	1343	0.006542	U
A-006	2900	3000	1260	0.007770	U
A-005	3200	2900	1222	0.009301	U
E-001	3700	1000	448	0.000027	L
E-004	3200	600	279	0.000046	L
E-002	3400	800	363	0.000054	L
E-003	3300	1200	526	0.000095	L
E-007	2800	700	316	0.000125	L
D-001	3000	1400	605	0.000136	L
E-005	3100	900	401	0.000189	L
E-006	3000	1200	523	0.000207	L
D-002	2500	1300	559	0.000251	L
D-004	2100	1300	555	0.000289	L
E-008	2700	1000	438	0.000304	L
D-003	2600	1700	724	0.000322	L
D-005	2200	1600	679	0.000379	L
D-007	1800	1500	634	0.000417	L
D-006	2300	1800	762	0.000543	L
D-008	1900	1800	758	0.000676	L

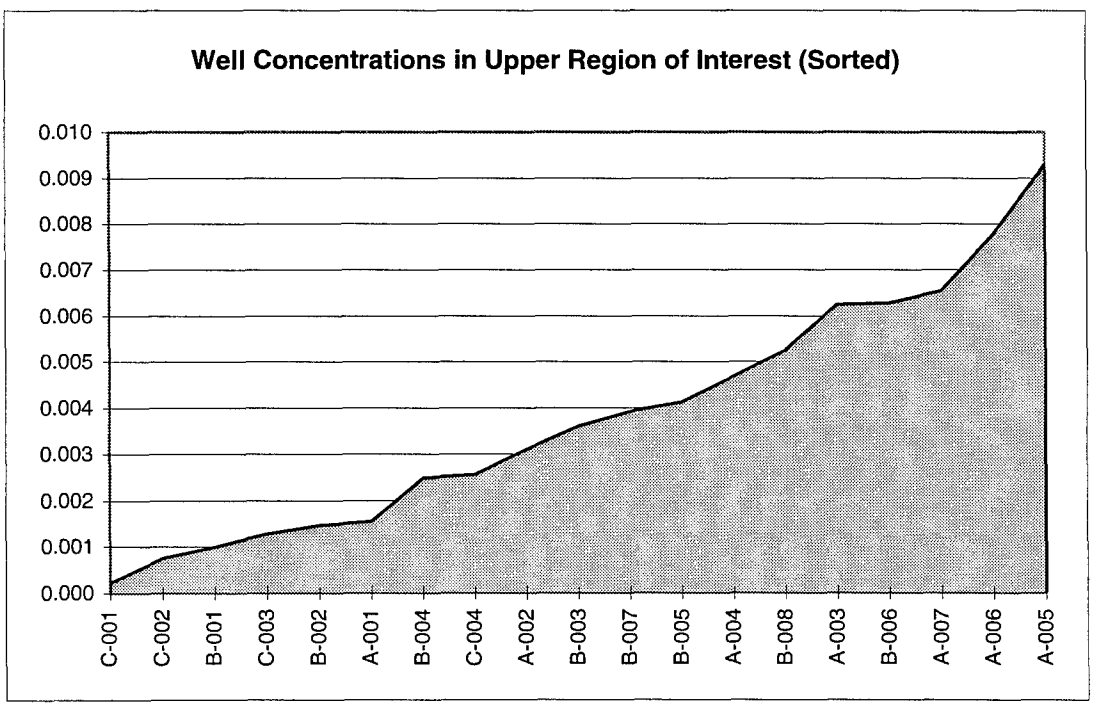
RSM Experiment One
Developing Regions of Possibility and Areas of Interest



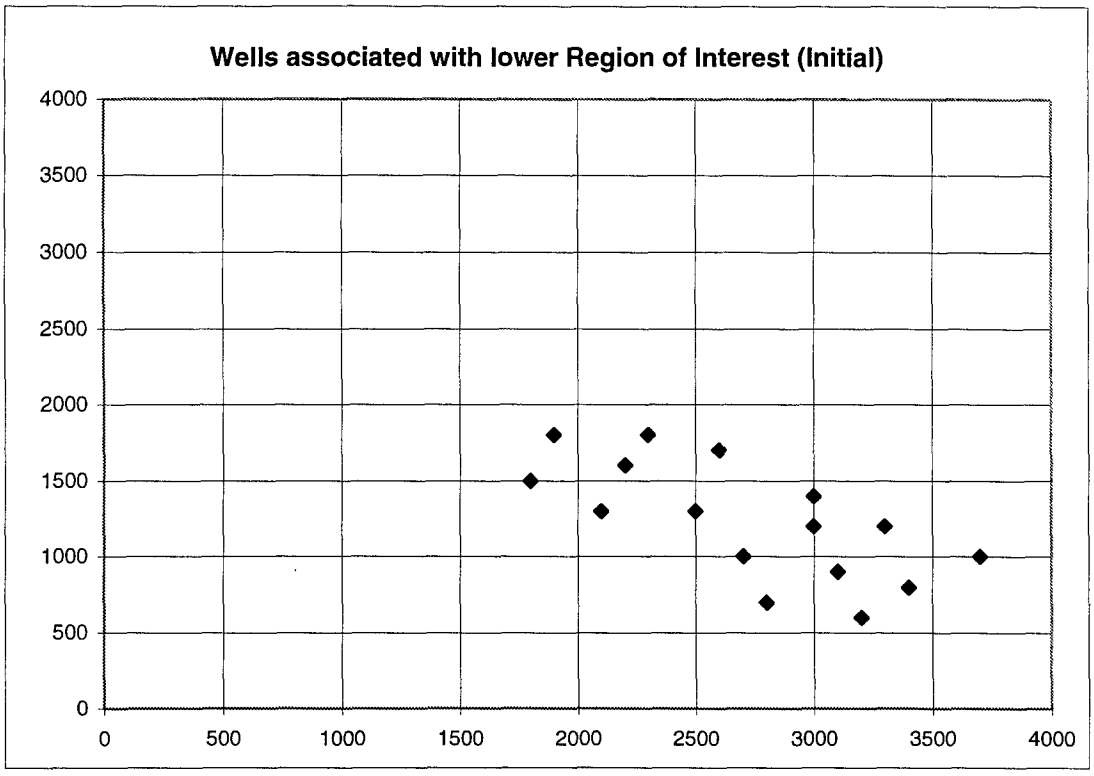


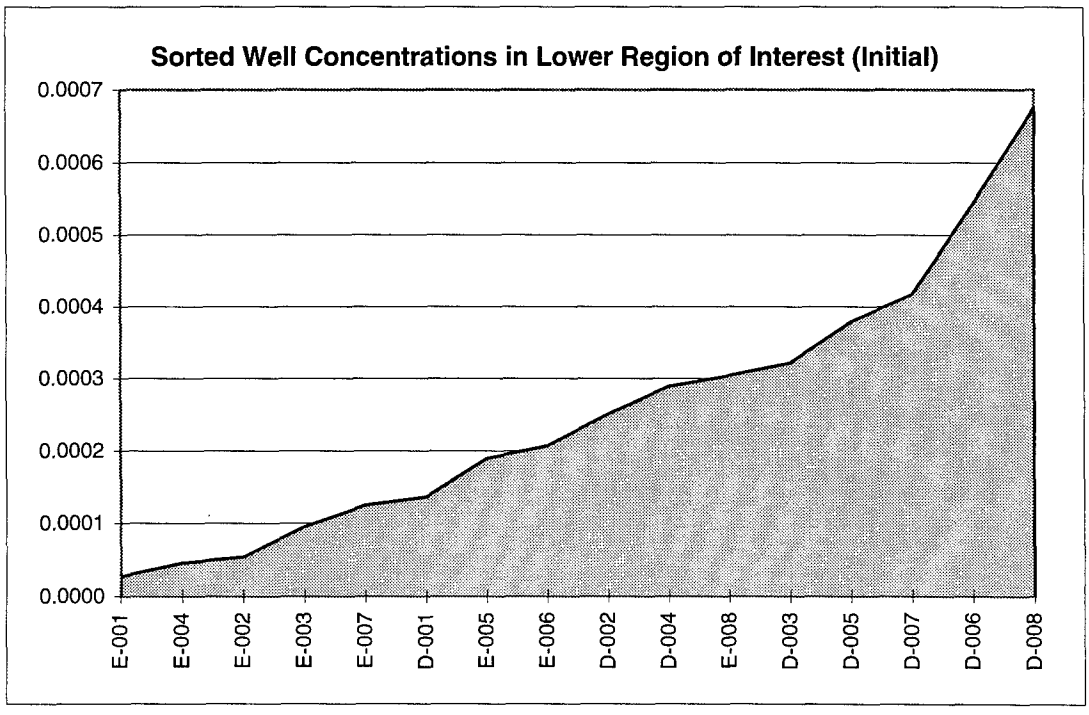
RSM Experiment One
Developing Regions of Possibility and Areas of Interest





RSM Experiment One
Developing Regions of Possibility and Areas of Interest





RSM Experiment One
Developing Regions of Possibility and Areas of Interest

Wells Sorted by Concentration
Separation of High and Low Concentration Sites

Well	X Loc.	Y Loc.	Node	Conc.	U, M or L
C-001	3300	1900	813	0.000220	U
C-002	2900	2100	891	0.000755	U
B-001	3200	2300	976	0.000993	U
C-003	2700	2200	930	0.001277	U
B-002	2900	2300	973	0.001465	U
A-001	3700	2700	1145	0.001548	U
B-004	2600	2400	1011	0.002480	U
C-004	2400	2400	1009	0.002560	U
A-002	3400	2700	1142	0.003093	U
B-003	3000	2600	1097	0.003598	U
B-007	2300	2700	1131	0.003924	U
B-005	2700	2600	1094	0.004125	U
A-004	3100	2700	1139	0.004686	U
B-008	2400	2900	1214	0.005253	U
A-003	3500	2900	1225	0.006247	U
B-006	2800	2800	1177	0.006285	U
A-007	3000	3200	1343	0.006542	U
A-006	2900	3000	1260	0.007770	U
A-005	3200	2900	1222	0.009301	U
D-003	2600	1700	724	0.000322	M
D-005	2200	1600	679	0.000379	M
D-007	1800	1500	634	0.000417	M
D-006	2300	1800	762	0.000543	M
D-008	1900	1800	758	0.000676	M
D-002	2500	1300	559	0.000251	M
D-004	2100	1300	555	0.000289	M
E-008	2700	1000	438	0.000304	M
E-001	3700	1000	448	0.000027	L
E-004	3200	600	279	0.000046	L
E-002	3400	800	363	0.000054	L
E-003	3300	1200	526	0.000095	L
E-007	2800	700	316	0.000125	L
D-001	3000	1400	605	0.000136	L
E-005	3100	900	401	0.000189	L
E-006	3000	1200	523	0.000207	L

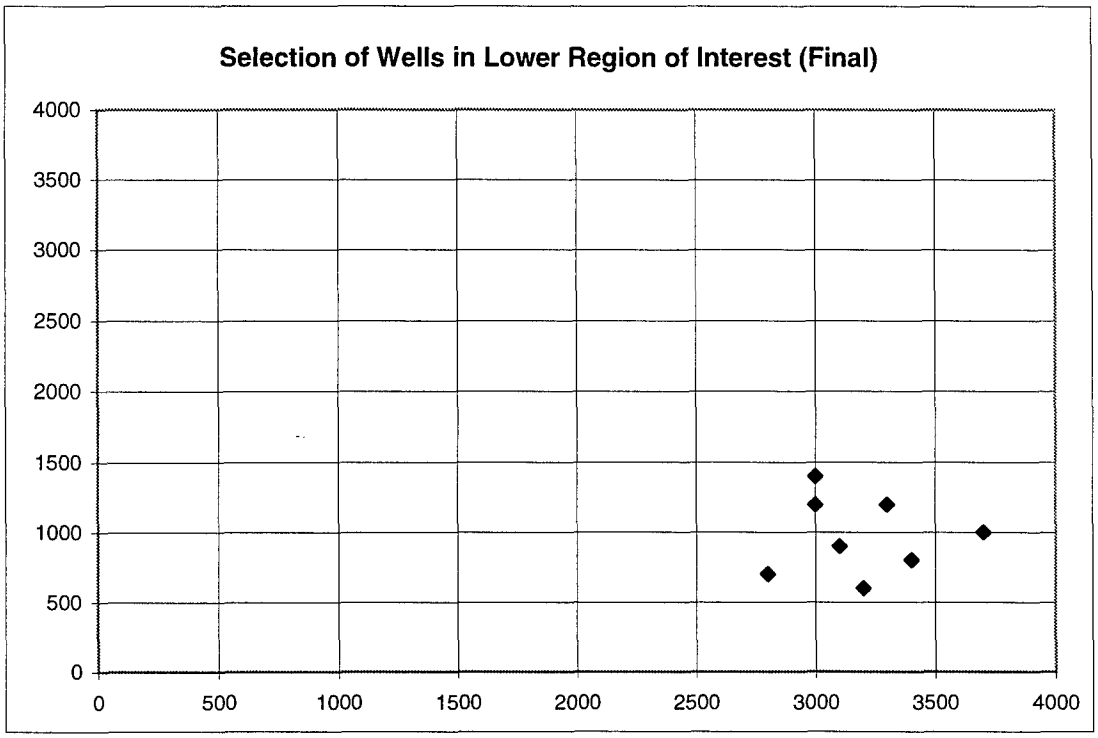
The addition of a middle zone came about when we realized the possibility of mixing exists where the two zones merged.

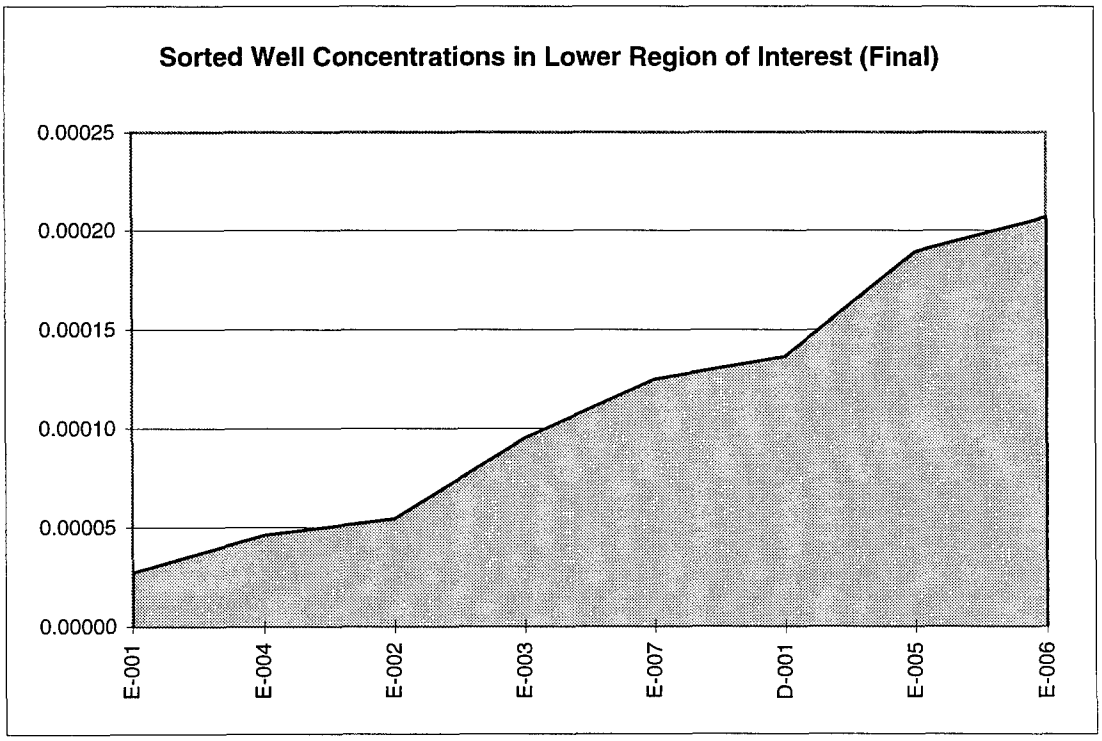
We examined the concentration levels in this middle zone and eliminated those measurement wells which we considered were unduly affected by either the upper and lower zone. The data to the left identify those wells which we considered to be so influenced.

Note that these middle zone wells, if considered along with the lower zone wells, would cause for an inability to "see" the concentration "hump" located in the middle zone, the key which identifies the existance of a source in this lower zone.

The existance of contaminant concentrations at well E-001 made clear the possibility of a source in the lower zone just by virtue of its location with respect to the historical sites of potential contaminant leakage/spillage and the groundwater flow lines.

RSM Experiment One
Developing Regions of Possibility and Areas of Interest





RSM Experiment One
Developing Regions of Possibility and Areas of Interest

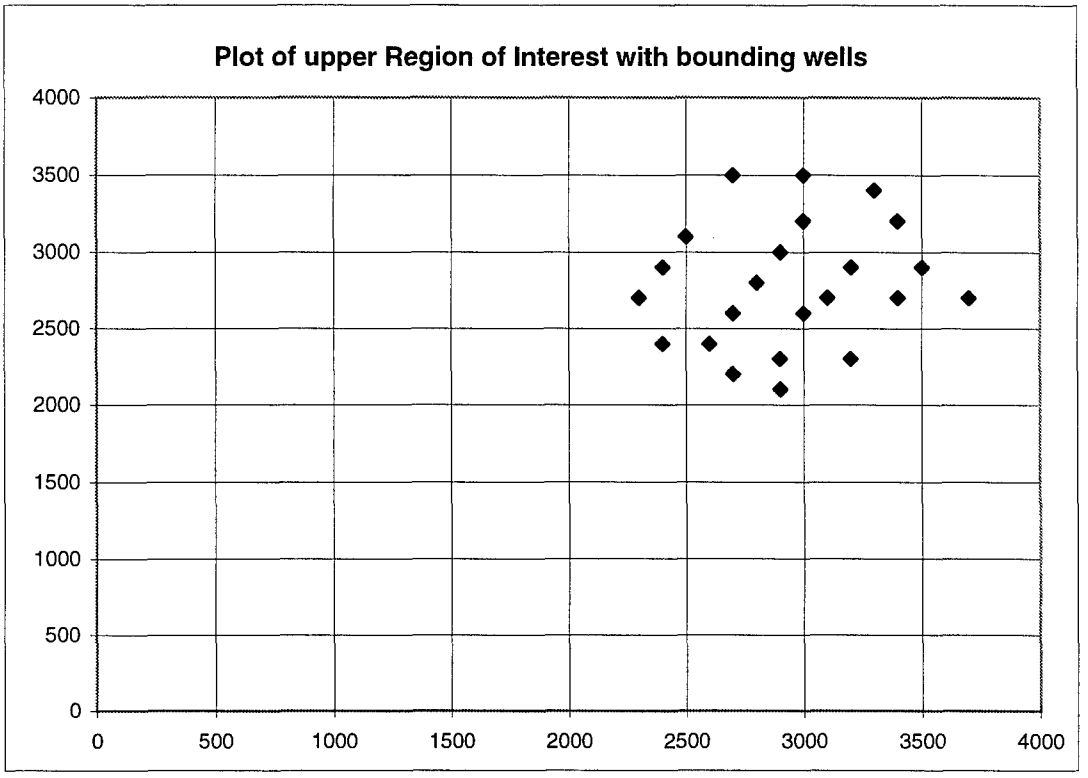
Wells Sorted by Concentration
Upper Contaminant Zone Wells Plus Additional Bounding Wells

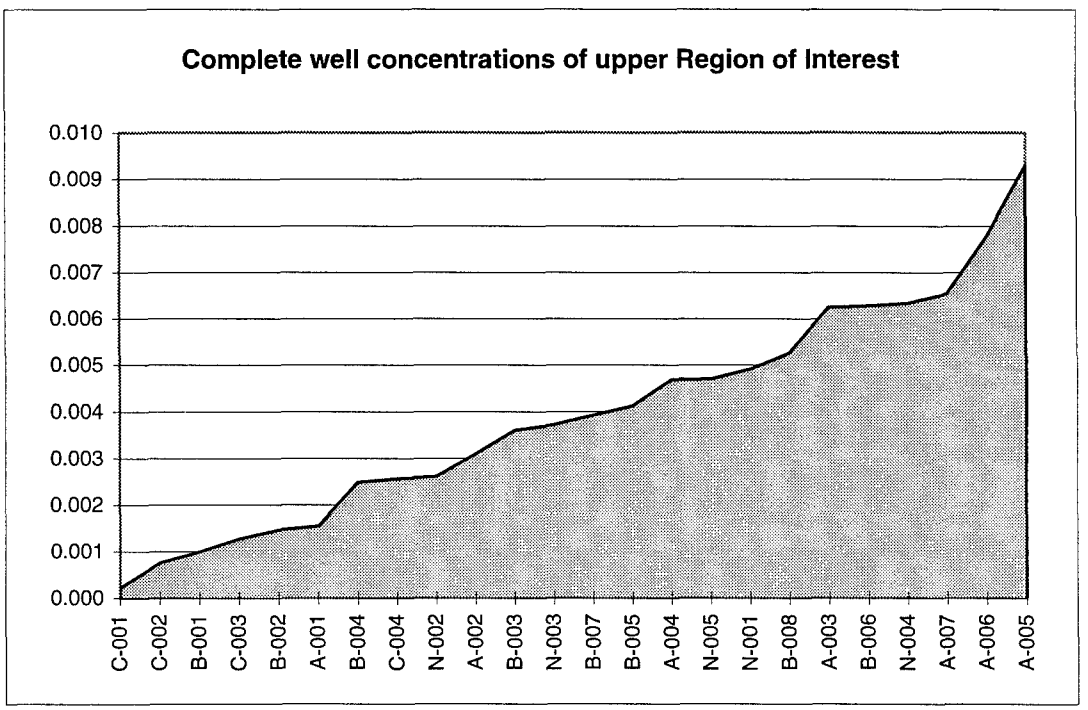
Well	X Loc.	Y Loc.	Node	Conc.
C-001	3300	1900	813	0.000220
C-002	2900	2100	891	0.000755
B-001	3200	2300	976	0.000993
C-003	2700	2200	930	0.001277
B-002	2900	2300	973	0.001465
A-001	3700	2700	1145	0.001548
B-004	2600	2400	1011	0.002480
C-004	2400	2400	1009	0.002560
N-002	3300	3400	1428	0.002604
A-002	3400	2700	1142	0.003093
B-003	3000	2600	1097	0.003598
N-003	3000	3500	1466	0.003715
B-007	2300	2700	1131	0.003924
B-005	2700	2600	1094	0.004125
A-004	3100	2700	1139	0.004686
N-005	2700	3500	1463	0.004689
N-001	3400	3200	1347	0.004910
B-008	2400	2900	1214	0.005253
A-003	3500	2900	1225	0.006247
B-006	2800	2800	1177	0.006285
N-004	2500	3100	1297	0.006341
A-007	3000	3200	1343	0.006542
A-006	2900	3000	1260	0.007770
A-005	3200	2900	1222	0.009301

The N-000 wells identified here are additional measurement wells which were placed in the upper and lower zones respectively.

They were used to further define the boundaries of the Region of Interest, as we looked for a pattern of higher concentration wells surrounded by lower concentration wells.

RSM Experiment One
Developing Regions of Possibility and Areas of Interest



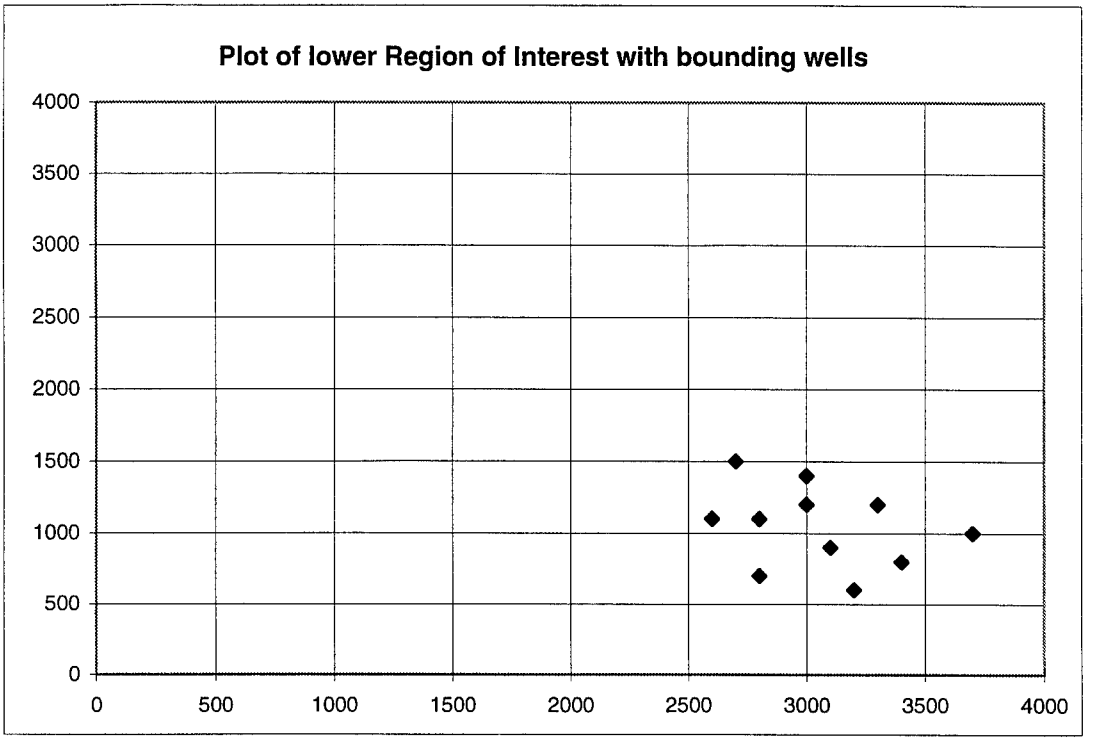


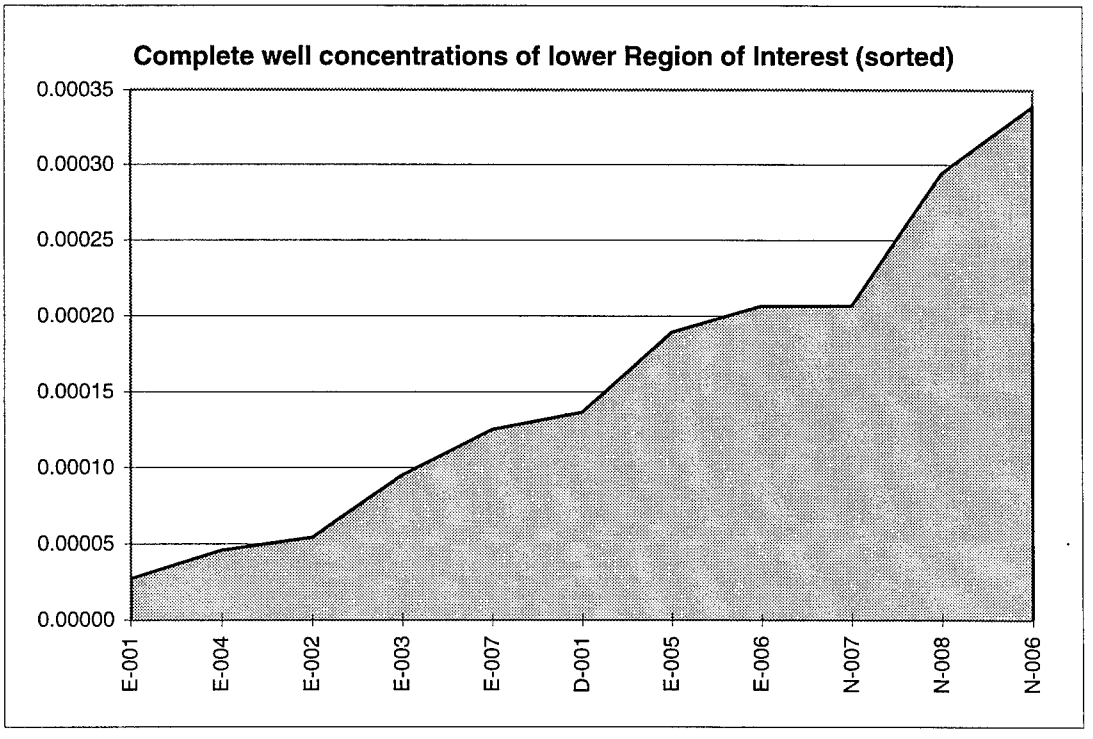
RSM Experiment One
Developing Regions of Possibility and Areas of Interest

Wells Sorted by Concentration
Lower Concentration Zone Plus Additional Bounding Wells

Well	X Loc.	Y Loc.	Node	Conc.
E-001	3700	1000	448	0.000027
E-004	3200	600	279	0.000046
E-002	3400	800	363	0.000054
E-003	3300	1200	526	0.000095
E-007	2800	700	316	0.000125
D-001	3000	1400	605	0.000136
E-005	3100	900	401	0.000189
E-006	3000	1200	523	0.000207
N-007	2700	1500	643	0.000207
N-008	2600	1100	478	0.000295
N-006	2800	1100	480	0.000339

RSM Experiment One
Developing Regions of Possibility and Areas of Interest





RSM Experiment One
Developing Regions of Possibility and Areas of Interest

Nodes To Be Used In:

Initial Screening

Node #	X Loc	Y Loc
355	2600	800
358	2900	800
436	2500	1000
439	2800	1000
442	3100	1000
445	3400	1000
516	2300	1200
519	2600	1200
522	2900	1200
525	3200	1200
600	2500	1400
1141	3300	2700
1144	3600	2700
1219	2900	2900
1222	3200	2900
1225	3500	2900
1299	2700	3100
1302	3000	3100
1305	3300	3100

Final Screening

Node #	X Loc	Y Loc
358	2900	800
399	2900	900
400	3000	900
440	2900	1000
441	3000	1000
481	2900	1100
482	3000	1100
483	3100	1100
523	3000	1200
524	3100	1200
525	3200	1200
1141	3300	2700
1182	3300	2800
1183	3400	2800
1223	3300	2900
1224	3400	2900
1264	3300	3000
1265	3400	3000
1305	3300	3100

The above nodes identify the nodes which are to be used in each of the two screenings for the experiment number one.

The wells to the right show the locations and concentrations obtained from the "truth set," which will be used to generate the sum of

Truth Values obtained From Measurement Wells

Well	Node	X Loc.	Y Loc.	Conc.
E-004	279	3200	600	0.000046
E-007	316	2800	700	0.000125
E-002	363	3400	800	0.000054
E-005	401	3100	900	0.000189
E-001	448	3700	1000	0.000027
N-008	478	2600	1100	0.000295
N-006	480	2800	1100	0.000339
E-006	523	3000	1200	0.000207
E-003	526	3300	1200	0.000095
D-001	605	3000	1400	0.000136
N-007	643	2700	1500	0.000207
C-001	813	3300	1900	0.000220
C-002	891	2900	2100	0.000755
C-003	930	2700	2200	0.001277
B-002	973	2900	2300	0.001465
B-001	976	3200	2300	0.000993
C-004	1009	2400	2400	0.002560
B-004	1011	2600	2400	0.002480
B-005	1094	2700	2600	0.004125
B-003	1097	3000	2600	0.003598
B-007	1131	2300	2700	0.003924
A-004	1139	3100	2700	0.004686
A-002	1142	3400	2700	0.003093
A-001	1145	3700	2700	0.001548
B-006	1177	2800	2800	0.006285
B-008	1214	2400	2900	0.005253
A-005	1222	3200	2900	0.009301
A-003	1225	3500	2900	0.006247
A-006	1260	2900	3000	0.007770
N-004	1297	2500	3100	0.006341
A-007	1343	3000	3200	0.006542
N-001	1347	3400	3200	0.004910
N-002	1428	3300	3400	0.002604
N-005	1463	2700	3500	0.004689
N-003	1466	3000	3500	0.003715

Flow Rate in upper Region of Interest ranges from 0.00000001 to 0.001 cu m/day
Flow Rate in lower Region of Interest ranges from 0.00000001 to 0.0001 cu m/day

#Wells: 19

Coded Values for Flow Rate -- Initial Screening

	Lower Concentration Zone											Upper Concentration Zone								
Runs	319	322	361	397	400	403	478	481	484	562	565	1142	1222	1225	1302	1304	1306	1384	1387	SSE Error
1	+	+	-	-	+	+	+	+	-	+	-	+	-	-	-	-	+	+	-	**0.002376**
2	-	+	+	-	-	+	+	+	+	-	+	-	+	-	-	-	-	+	+	**0.003431**
3	+	-	+	+	-	-	+	+	+	+	-	+	-	+	-	-	-	-	+	**0.004202**
4	+	+	-	+	+	-	-	+	+	+	+	-	+	-	+	-	-	-	-	**0.003303**
5	-	+	+	-	+	+	-	-	+	+	+	+	-	+	-	+	-	-	-	**0.006206**
6	-	-	+	+	-	+	+	-	-	+	+	+	+	-	+	-	+	-	-	**0.010905**
7	-	-	-	+	+	-	+	+	-	-	+	+	+	+	-	+	-	+	-	**0.016791**
8	-	-	-	-	+	+	-	+	+	-	-	+	+	+	+	-	+	-	+	**0.018577**
9	+	-	-	-	-	+	+	-	+	+	-	-	+	+	+	+	-	+	-	**0.012823**
10	-	+	-	-	-	-	+	+	-	+	+	-	-	+	+	+	+	-	+	**0.007306**
11	+	-	+	-	-	-	-	+	+	-	+	+	-	-	+	+	+	+	-	**0.010591**
12	-	+	-	+	-	-	-	-	+	+	-	+	+	-	-	+	+	+	+	**0.018271**
13	+	-	+	-	+	-	-	-	-	+	+	-	+	+	-	-	+	+	+	**0.006626**
14	+	+	-	+	-	+	-	-	-	-	+	+	-	+	+	-	-	+	+	**0.009921**
15	+	+	+	-	+	-	+	-	-	-	-	+	+	-	+	+	-	-	+	**0.020231**
16	+	+	+	+	-	+	-	+	-	-	-	-	+	+	-	+	+	-	-	**0.007347**
17	-	+	+	+	+	-	+	-	+	-	-	-	-	+	+	-	+	+	-	**0.002601**
18	-	-	+	+	+	+	-	+	-	+	-	-	-	-	+	+	-	+	+	**0.005076**
19	+	-	-	+	+	+	+	-	+	-	+	-	-	-	-	+	+	-	+	**0.001845**
20	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	**0.000097**

LEAST SQUARES SOLUTION (MathCad)

$$\text{Solution} := \left(A^{T} \cdot A\right)^{-1} \cdot A^{T} \cdot Z$$

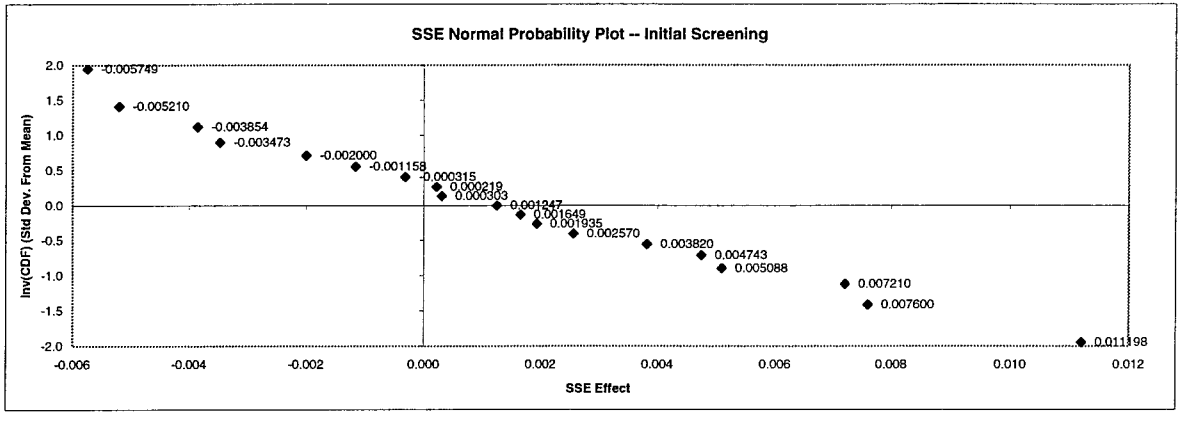
$$A = \begin{bmatrix}
1 & 1 & -1 & -1 & 1 & 1 & 1 & 1 & -1 & 1 & -1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 & -1 \\
-1 & 1 & 1 & -1 & -1 & 1 & 1 & 1 & 1 & -1 & 1 & -1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 \\
1 & -1 & 1 & 1 & -1 & -1 & 1 & 1 & 1 & 1 & -1 & -1 & -1 & 1 & -1 & -1 & -1 & -1 & 1 \\
1 & 1 & -1 & 1 & 1 & -1 & -1 & 1 & 1 & 1 & 1 & 1 & -1 & 1 & -1 & -1 & -1 & -1 & -1 \\
-1 & 1 & 1 & -1 & 1 & 1 & -1 & -1 & 1 & 1 & 1 & 1 & -1 & 1 & -1 & 1 & -1 & -1 & -1 \\
-1 & -1 & 1 & 1 & -1 & 1 & 1 & -1 & -1 & 1 & 1 & 1 & 1 & -1 & 1 & -1 & 1 & -1 & -1 \\
-1 & -1 & -1 & 1 & 1 & -1 & 1 & 1 & -1 & -1 & 1 & -1 & 1 & 1 & -1 & 1 & -1 & 1 & -1 \\
-1 & -1 & -1 & -1 & 1 & 1 & -1 & 1 & 1 & -1 & -1 & 1 & 1 & 1 & -1 & 1 & -1 & -1 & 1 \\
1 & -1 & -1 & -1 & -1 & 1 & 1 & -1 & 1 & 1 & -1 & 1 & 1 & 1 & 1 & 1 & -1 & 1 & -1 \\
-1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 & -1 & 1 & 1 & -1 & 1 & 1 & 1 & 1 & -1 & 1 & 1 \\
1 & -1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 & -1 & 1 & 1 & -1 & 1 & 1 & 1 & 1 & 1 & -1 \\
-1 & 1 & -1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 & -1 & 1 & 1 & 1 & -1 & -1 & 1 & 1 & 1 \\
1 & 1 & -1 & 1 & -1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 & -1 & 1 & 1 & -1 & -1 & 1 & 1 \\
1 & 1 & 1 & -1 & 1 & -1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 & -1 & 1 & 1 & -1 & -1 & 1 \\
1 & 1 & 1 & 1 & -1 & 1 & -1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 & -1 & 1 & 1 & -1 & -1 \\
-1 & 1 & 1 & 1 & 1 & -1 & 1 & -1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 & -1 & 1 & 1 & -1 \\
-1 & -1 & 1 & 1 & 1 & 1 & -1 & 1 & -1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 & -1 & 1 & 1 \\
1 & -1 & -1 & 1 & 1 & 1 & 1 & -1 & 1 & -1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 & -1 & 1 \\
-1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1
\end{bmatrix}$$

$$Z := \begin{bmatrix}
0.009608 \\ 0.013119 \\ 0.012702 \\ 0.007639 \\ 0.016101 \\ 0.024394 \\ 0.042798 \\ 0.056168 \\ 0.046345 \\ 0.042713 \\ 0.040913 \\ 0.056839 \\ 0.033755 \\ 0.042373 \\ 0.051828 \\ 0.023461 \\ 0.023029 \\ 0.032539 \\ 0.013224 \\ 0.000502
\end{bmatrix}$$

$$\text{Solution} = \begin{bmatrix}
-0.00425265 \\ -0.0067014 \\ 6.1655 \cdot 10^{-4} \\ 0.00133225 \\ 0.00210135 \\ -0.0017693 \\ -0.0015265 \\ 0.00159845 \\ -8.946 \cdot 10^{-4} \\ -0.0071089 \\ -0.00473455 \\ 0.01467475 \\ 0.0061321 \\ 0.00737695 \\ 0.0014217 \\ 0.00423865 \\ 0.0029079 \\ 0.0046293 \\ 0.00895845
\end{bmatrix}$$

Equivalent Actual Values for Flow Rate -- Initial Screening

	Lower Concentration Zone											Upper Concentration Zone								
	319	322	361	397	400	403	478	481	484	562	565	1142	1222	1225	1302	1304	1306	1384	1387	SSE Error
1	1.00E-04	1.00E-04	1.00E-08	1.00E-08	1.00E-04	1.00E-04	1.00E-04	1.00E-04	1.00E-08	1.00E-04	1.00E-08	1.00E-03	1.00E-08	1.00E-08	1.00E-08	1.00E-08	1.00E-03	1.00E-03	1.00E-08	**0.005957**
2	1.00E-08	1.00E-04	1.00E-04	1.00E-08	1.00E-08	1.00E-04	1.00E-04	1.00E-04	1.00E-04	1.00E-08	1.00E-04	1.00E-08	1.00E-03	1.00E-08	1.00E-08	1.00E-08	1.00E-08	1.00E-03	1.00E-03	**0.005554**
3	1.00E-04	1.00E-08	1.00E-04	1.00E-04	1.00E-08	1.00E-08	1.00E-04	1.00E-04	1.00E-04	1.00E-04	1.00E-08	1.00E-03	1.00E-08	1.00E-03	1.00E-08	1.00E-08	1.00E-08	1.00E-08	1.00E-03	**0.007155**
4	1.00E-04	1.00E-04	1.00E-08	1.00E-04	1.00E-04	1.00E-08	1.00E-08	1.00E-04	1.00E-04	1.00E-04	1.00E-04	1.00E-08	1.00E-03	1.00E-08	1.00E-03	1.00E-08	1.00E-08	1.00E-08	1.00E-08	**0.002628**
5	1.00E-08	1.00E-04	1.00E-04	1.00E-08	1.00E-04	1.00E-04	1.00E-08	1.00E-08	1.00E-04	1.00E-04	1.00E-04	1.00E-03	1.00E-08	1.00E-03	1.00E-08	1.00E-03	1.00E-08	1.00E-08	1.00E-08	**0.007654**
6	1.00E-08	1.00E-08	1.00E-04	1.00E-04	1.00E-08	1.00E-04	1.00E-04	1.00E-08	1.00E-08	1.00E-04	1.00E-04	1.00E-03	1.00E-03	1.00E-08	1.00E-03	1.00E-08	1.00E-03	1.00E-08	1.00E-08	**0.011195**
7	1.00E-08	1.00E-08	1.00E-08	1.00E-04	1.00E-04	1.00E-08	1.00E-04	1.00E-04	1.00E-08	1.00E-08	1.00E-04	1.00E-03	1.00E-03	1.00E-03	1.00E-08	1.00E-03	1.00E-08	1.00E-03	1.00E-08	**0.018330**
8	1.00E-08	1.00E-08	1.00E-08	1.00E-08	1.00E-04	1.00E-04	1.00E-08	1.00E-04	1.00E-04	1.00E-08	1.00E-08	1.00E-03	1.00E-03	1.00E-03	1.00E-03	1.00E-08	1.00E-03	1.00E-08	1.00E-03	**0.027209**
9	1.00E-04	1.00E-08	1.00E-08	1.00E-08	1.00E-08	1.00E-04	1.00E-04	1.00E-08	1.00E-04	1.00E-04	1.00E-08	1.00E-08	1.00E-03	1.00E-03	1.00E-03	1.00E-03	1.00E-08	1.00E-03	1.00E-08	**0.017851**
10	1.00E-08	1.00E-04	1.00E-08	1.00E-08	1.00E-08	1.00E-08	1.00E-04	1.00E-04	1.00E-08	1.00E-04	1.00E-04	1.00E-08	1.00E-08	1.00E-03	1.00E-03	1.00E-03	1.00E-03	1.00E-08	1.00E-03	**0.019627**
11	1.00E-04	1.00E-08	1.00E-04	1.00E-08	1.00E-08	1.00E-08	1.00E-08	1.00E-04	1.00E-04	1.00E-08	1.00E-04	1.00E-03	1.00E-08	1.00E-08	1.00E-03	1.00E-03	1.00E-03	1.00E-03	1.00E-08	**0.017731**
12	1.00E-08	1.00E-04	1.00E-08	1.00E-04	1.00E-08	1.00E-08	1.00E-08	1.00E-08	1.00E-04	1.00E-04	1.00E-08	1.00E-03	1.00E-03	1.00E-08	1.00E-08	1.00E-03	1.00E-03	1.00E-03	1.00E-03	**0.024840**
13	1.00E-04	1.00E-08	1.00E-04	1.00E-08	1.00E-04	1.00E-08	1.00E-08	1.00E-08	1.00E-08	1.00E-04	1.00E-04	1.00E-08	1.00E-03	1.00E-03	1.00E-08	1.00E-08	1.00E-03	1.00E-03	1.00E-03	**0.018114**
14	1.00E-04	1.00E-04	1.00E-08	1.00E-04	1.00E-08	1.00E-04	1.00E-08	1.00E-08	1.00E-08	1.00E-08	1.00E-04	1.00E-03	1.00E-08	1.00E-03	1.00E-03	1.00E-08	1.00E-08	1.00E-03	1.00E-03	**0.017298**
15	1.00E-04	1.00E-04	1.00E-04	1.00E-08	1.00E-04	1.00E-08	1.00E-04	1.00E-08	1.00E-08	1.00E-08	1.00E-08	1.00E-03	1.00E-03	1.00E-08	1.00E-03	1.00E-03	1.00E-08	1.00E-08	1.00E-03	**0.016528**
16	1.00E-04	1.00E-04	1.00E-04	1.00E-04	1.00E-08	1.00E-04	1.00E-08	1.00E-04	1.00E-08	1.00E-08	1.00E-08	1.00E-08	1.00E-03	1.00E-03	1.00E-08	1.00E-03	1.00E-03	1.00E-08	1.00E-08	**0.012963**
17	1.00E-08	1.00E-04	1.00E-04	1.00E-04	1.00E-04	1.00E-08	1.00E-04	1.00E-08	1.00E-04	1.00E-08	1.00E-08	1.00E-08	1.00E-08	1.00E-03	1.00E-03	1.00E-08	1.00E-03	1.00E-03	1.00E-08	**0.012102**
18	1.00E-08	1.00E-08	1.00E-04	1.00E-04	1.00E-04	1.00E-04	1.00E-08	1.00E-04	1.00E-08	1.00E-04	1.00E-08	1.00E-08	1.00E-08	1.00E-08	1.00E-03	1.00E-03	1.00E-08	1.00E-03	1.00E-03	**0.012823**
19	1.00E-04	1.00E-08	1.00E-08	1.00E-04	1.00E-04	1.00E-04	1.00E-04	1.00E-08	1.00E-04	1.00E-08	1.00E-04	1.00E-08	1.00E-08	1.00E-08	1.00E-08	1.00E-03	1.00E-03	1.00E-08	1.00E-03	**0.007203**
20	1.00E-08	1.00E-08	1.00E-08	1.00E-08	1.00E-08	1.00E-08	1.00E-08	1.00E-08	1.00E-08	1.00E-08	1.00E-08	1.00E-08	1.00E-08	1.00E-08	1.00E-08	1.00E-08	1.00E-08	1.00E-08	1.00E-08	**0.000239**
Coefficient	-0.001927	-0.002875	0.000152	0.000624	0.000825	-0.000579	-0.001000	0.000967	-0.000157	-0.002605	-0.001736	0.005599	0.002371	0.003800	0.000110	0.001285	0.002544	0.001910	0.003605	
EFFECT	-0.003854	-0.005749	0.000303	0.001247	0.001649	-0.001158	-0.002000	0.001935	-0.000315	-0.005210	-0.003473	0.011198	0.004743	0.007600	0.000219	0.002570	0.005088	0.003820	0.007210	
Rank	17	19	11	10	9	14	15	8	13	18	16	1	5	2	12	7	4	6	3	
CDF	0.8684	0.9737	0.5526	0.5000	0.4474	0.7105	0.7632	0.3947	0.6579	0.9211	0.8158	0.0263	0.2368	0.0789	0.6053	0.3421	0.1842	0.2895	0.1316	
Step 1	0.9811	0.9964	0.9231	0.9107	0.8972	0.9549	0.9642	0.8821	0.9451	0.9890	0.9729	0.6122	0.8234	0.7100	0.9345	0.8653	0.7960	0.8460	0.7606	
Step 2	0.1316	0.0263	0.4474	0.5000	0.5526	0.2895	0.2368	0.6053	0.3421	0.0789	0.1842	0.9737	0.7632	0.9211	0.3947	0.6579	0.8158	0.7105	0.8684	
Step 3	0.7606	0.6122	0.8972	0.9107	0.9231	0.8460	0.8234	0.9345	0.8653	0.7100	0.7960	0.9964	0.9642	0.9890	0.8821	0.9451	0.9729	0.9549	0.9811	
Inv(CDF)	1.1165	1.9454	0.1314	0.0000	-0.1314	0.5516	0.7128	-0.2651	0.4040	1.4126	0.8959	-1.9454	-0.7128	-1.4126	0.2651	-0.4040	-0.8959	-0.5516	-1.1165	



Flow Rate in upper Region of Interest ranges from 0.00000001 to 0.001 cu m/day
Flow Rate in lower Region of Interest ranges from 0.00000001 to 0.0001 cu m/day

#Wells: 19 Equivalent Actual Values for Flow Rate -- Final Screening

	Lower Concentration Zone						Upper Concentration Zone					
	280	321	322	362	363	1141	1142	1143	1182	1183	1184	SSE ERROR
1	+	+	-	+	+	+	-	-	-	+	-	**0.0013789944**
2	-	+	+	-	+	+	+	-	-	-	+	**0.0020604671**
3	+	-	+	+	-	+	+	+	-	-	-	**0.0017522005**
4	-	+	-	+	+	-	+	+	+	-	-	**0.0019792073**
5	-	-	+	-	+	+	-	+	+	+	-	**0.0027507449**
6	-	-	-	+	-	+	+	-	+	+	+	**0.0036171004**
7	+	-	-	-	+	-	+	+	-	+	+	**0.0030149023**
8	+	+	-	-	-	+	-	+	+	-	+	**0.0028630007**
9	+	+	+	-	-	-	+	-	+	+	-	**0.0021753805**
10	-	+	+	+	-	-	-	+	-	+	+	**0.0024366575**
11	+	-	+	+	+	-	-	-	+	-	+	**0.0017029848**
12	-	-	-	-	-	-	-	-	-	-	-	**0.0000000441**

LEAST SQUARES SOLUTION (MathCad)

$$\text{Solution} := \left(A^{T} \cdot A\right)^{-1} \cdot A^{T} \cdot Z$$

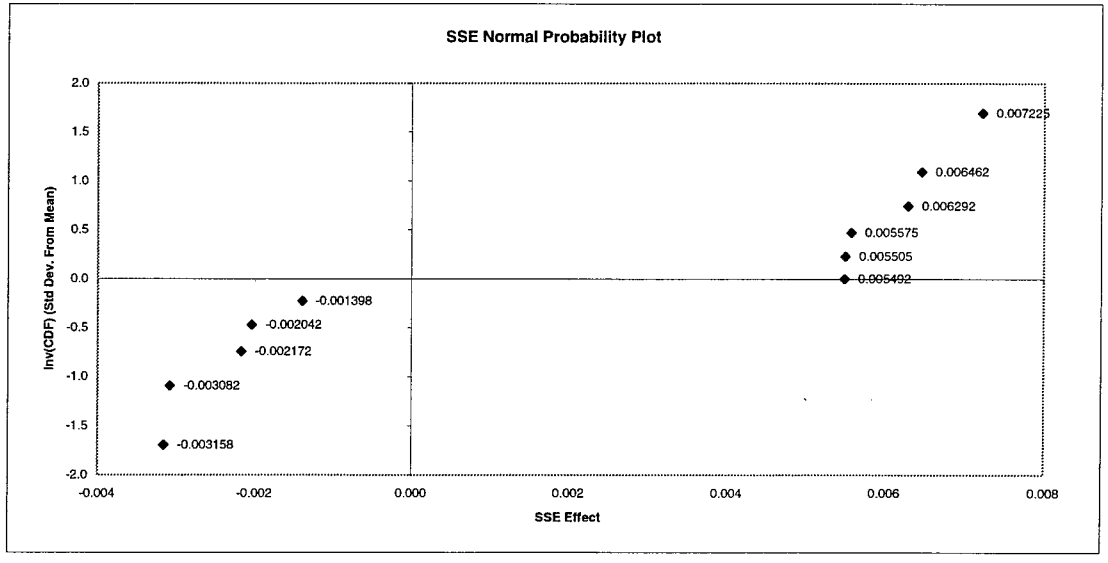
$$A = \begin{bmatrix}
1 & 1 & -1 & -1 & 1 & 1 & 1 & 1 & -1 & 1 & -1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 & -1 \\
-1 & 1 & 1 & -1 & -1 & 1 & 1 & 1 & 1 & -1 & 1 & -1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 \\
1 & -1 & 1 & 1 & -1 & -1 & 1 & 1 & 1 & 1 & -1 & -1 & 1 & -1 & -1 & -1 & -1 & 1 & \\
1 & 1 & -1 & 1 & 1 & -1 & -1 & 1 & 1 & 1 & 1 & 1 & -1 & 1 & -1 & -1 & -1 & -1 & \\
-1 & 1 & 1 & -1 & 1 & 1 & -1 & -1 & 1 & 1 & 1 & 1 & -1 & 1 & -1 & 1 & -1 & -1 & \\
-1 & -1 & 1 & 1 & -1 & 1 & 1 & -1 & -1 & 1 & 1 & 1 & 1 & -1 & 1 & -1 & 1 & -1 & \\
-1 & -1 & -1 & 1 & 1 & -1 & 1 & 1 & -1 & -1 & 1 & -1 & 1 & 1 & -1 & 1 & -1 & 1 & \\
-1 & -1 & -1 & -1 & 1 & 1 & -1 & 1 & 1 & -1 & -1 & -1 & 1 & 1 & 1 & -1 & 1 & -1 & \\
1 & -1 & -1 & -1 & -1 & 1 & 1 & -1 & 1 & 1 & -1 & 1 & 1 & 1 & 1 & 1 & -1 & 1 & \\
-1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 & -1 & 1 & 1 & 1 & -1 & 1 & 1 & 1 & 1 & -1 & \\
1 & -1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 & -1 & 1 & 1 & -1 & -1 & 1 & 1 & 1 & 1 & \\
-1 & 1 & -1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 & -1 & 1 & 1 & -1 & -1 & 1 & 1 & 1 & \\
1 & -1 & 1 & -1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 & -1 & 1 & 1 & -1 & -1 & 1 & 1 & \\
1 & 1 & -1 & 1 & -1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 & -1 & 1 & 1 & -1 & -1 & 1 & \\
1 & 1 & 1 & -1 & 1 & -1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 & -1 & 1 & 1 & -1 & -1 & \\
1 & 1 & 1 & 1 & -1 & 1 & -1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 & -1 & 1 & 1 & -1 & \\
-1 & 1 & 1 & 1 & 1 & -1 & 1 & -1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 & -1 & 1 & 1 & \\
-1 & -1 & 1 & 1 & 1 & 1 & -1 & 1 & -1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 & -1 & 1 & \\
1 & -1 & -1 & 1 & 1 & 1 & 1 & -1 & 1 & -1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 & -1 & 1 \\
-1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1
\end{bmatrix}$$

$$Z := \begin{bmatrix}
0.015332 \\ 0.015617 \\ 0.014916 \\ 0.007276 \\ 0.017536 \\ 0.030607 \\ 0.050788 \\ 0.070153 \\ 0.052821 \\ 0.050474 \\ 0.050572 \\ 0.070441 \\ 0.047947 \\ 0.04765 \\ 0.048603 \\ 0.031698 \\ 0.03178 \\ 0.032002 \\ 0.016531 \\ 0.000502
\end{bmatrix}$$

$$\text{Solution} = \begin{bmatrix}
-0.00507645 \\ -0.0080191 \\ 0.00021425 \\ 0.00145535 \\ 0.00188125 \\ -0.0021676 \\ -0.0024154 \\ 0.00196925 \\ -0.000398 \\ -0.0077246 \\ -0.00491125 \\ 0.01624375 \\ 0.0074328 \\ 0.00966275 \\ 0.000534 \\ 0.00373555 \\ 0.0063912 \\ 0.0063327 \\ 0.00951985
\end{bmatrix}$$

Equivalent Actual Values for Flow Rate -- Initial Screening

	Lower Concentration Zone					Upper Concentration Zone						
	280	321	322	362	363	1141	1142	1143	1182	1183	1184	SSE ERROR
1	1.00E-04	1.00E-04	1.00E-08	1.00E-04	1.00E-04	1.00E-03	1.00E-08	1.00E-08	1.00E-08	1.00E-03	1.00E-08	**0.00137899**
2	1.00E-08	1.00E-04	1.00E-04	1.00E-08	1.00E-04	1.00E-03	1.00E-03	1.00E-08	1.00E-08	1.00E-08	1.00E-03	**0.00206047**
3	1.00E-04	1.00E-08	1.00E-04	1.00E-04	1.00E-08	1.00E-03	1.00E-03	1.00E-03	1.00E-08	1.00E-08	1.00E-08	**0.00175220**
4	1.00E-08	1.00E-04	1.00E-08	1.00E-04	1.00E-04	1.00E-08	1.00E-03	1.00E-03	1.00E-03	1.00E-08	1.00E-08	**0.00197921**
5	1.00E-08	1.00E-08	1.00E-04	1.00E-08	1.00E-04	1.00E-03	1.00E-08	1.00E-03	1.00E-03	1.00E-03	1.00E-08	**0.00275074**
6	1.00E-08	1.00E-08	1.00E-08	1.00E-04	1.00E-08	1.00E-03	1.00E-03	1.00E-08	1.00E-03	1.00E-03	1.00E-03	**0.00361710**
7	1.00E-04	1.00E-08	1.00E-08	1.00E-08	1.00E-04	1.00E-08	1.00E-03	1.00E-03	1.00E-08	1.00E-03	1.00E-03	**0.00301490**
8	1.00E-04	1.00E-04	1.00E-08	1.00E-08	1.00E-08	1.00E-03	1.00E-08	1.00E-03	1.00E-03	1.00E-08	1.00E-03	**0.00286300**
9	1.00E-04	1.00E-04	1.00E-04	1.00E-08	1.00E-08	1.00E-08	1.00E-03	1.00E-08	1.00E-03	1.00E-03	1.00E-08	**0.00217538**
10	1.00E-08	1.00E-04	1.00E-04	1.00E-04	1.00E-08	1.00E-08	1.00E-08	1.00E-03	1.00E-08	1.00E-03	1.00E-03	**0.00243666**
11	1.00E-04	1.00E-08	1.00E-04	1.00E-04	1.00E-04	1.00E-08	1.00E-08	1.00E-08	1.00E-03	1.00E-08	1.00E-03	**0.00170298**
12	1.00E-08	1.00E-08	1.00E-08	1.00E-08	1.00E-08	1.00E-08	1.00E-08	1.00E-08	1.00E-08	1.00E-08	1.00E-08	**0.00000004**
Coefficient	-0.001021	-0.001541	-0.001579	-6.99E-04	-0.001086	0.0031458	0.0027525	0.00274583	0.002788	0.003231	0.003613	
EFFECT	-0.002042	-0.003082	-0.003158	-0.001398	-0.002172	0.006292	0.005505	0.005492	0.005575	0.006462	0.007225	
Rank	4	2	1	5	3	9	7	6	8	10	11	
CDF	0.3182	0.1364	0.0455	0.4091	0.2273	0.7727	0.5909	0.5000	0.6818	0.8636	0.9545	
Step 1	0.8569	0.7643	0.6590	0.8864	0.8188	0.9658	0.9315	0.9107	0.9496	0.9804	0.9937	
Step 2	0.6818	0.8636	0.9545	0.5909	0.7727	0.2273	0.4091	0.5000	0.3182	0.1364	0.0455	
Step 3	0.9496	0.9804	0.9937	0.9315	0.9658	0.8188	0.8864	0.9107	0.8569	0.7643	0.6590	
Inv(CDF)	-0.4698	-1.0942	-1.6947	-0.2283	-0.7442	0.7442	0.2283	0.0000	0.4698	1.0942	1.6947	



Design A Experimental Values for Experiment One
Calculating Flow Rates

Design A-1 Experimental Values:

Node Number	Low Flow Rate	Med Flow Rate	High Flow Rate
524	1.00E-08	5.00E-04	1.00E-03
1141	1.00E-08	5.00E-05	1.00E-04

Design A-1 Setup:

	Values to be used in 2^2 Design	
Run	Node 524	Node 1141
1	4.00E-05	4.00E-04
2	6.00E-05	4.00E-04
3	4.00E-05	6.00E-04
4	6.00E-05	6.00E-04
Center	5.00E-05	5.00E-04

SSE Values obtained from previous 2x2 run.

	Coded Values		**Design A-1**
Run	Node 524	Node 1141	SSE
1	-1	-1	0.0001825507
2	1	-1	0.0001858608
3	-1	1	0.0003282440
4	1	1	0.0003322036
Center	0	0	0.0002317981

Calculated Sum of Squares Solution obtained from SSE runs yields the reduction gradient (difference between run #1 and #2) and successive model run parameters.

Run Number	Low Zone Val	High Zone Val	SSE
1	0.00005	0.0005	0.0002317981
2	0.00004975	0.00040003	0.000184001
3	0.0000495	0.00030006	0.000186913
4	0.00004925	0.00020009	0.000240775
5	0.000049	0.00010012	0.000345827
6	0.00004876	0.00000015	0.000502313

Design A Experimental Values for Experiment One
Calculating Flow Rates

Design A-2 Experimental Values:

Node Number	Low Flow Rate	Med Flow Rate	High Flow Rate
524	0.000049500	0.000049625	0.000049750
1141	0.000400030	0.000350045	0.000300060

Design A-2 Setup:

	Values to be used in 2^2 Design	
Run	Node 524	Node 1141
1	0.000049500	0.000400030
2	0.000049750	0.000400030
3	0.000049500	0.000300060
4	0.000049750	0.000300060
Center	0.000049625	0.000350045

SSE Values obtained from previous 2x2 run.

	Coded Values		**Design A-2**
Run	Node 524	Node 1141	SSE
1	-1	-1	0.0001869133
2	1	-1	0.0001869503
3	-1	1	0.0001839596
4	1	1	0.0001840007
Center	0	0	0.0001791032

Calculated Sum of Squares Solution obtained from SSE runs yields the reduction gradient (difference between run #1 and #2) and successive model run parameters.

Run Number	Low Zone Val	High Zone Val	SSE
1	0.0000496250	0.000350045	0.000179103
2	0.0000496233	0.000400026	0.000183979
3	0.0000496217	0.000450006	0.000201528
4	0.0000496200	0.000499987	0.00023172
5	0.0000496184	0.000549968	0.000274525

Design A Experimental Values for Experiment One
Calculating Flow Rates

Design A-3 Experimental Values:

Node Number	Low Flow Rate	Med Flow Rate	High Flow Rate
524	0.0000496233	0.0000496242	0.0000496250
1141	0.0003500450	0.0003750353	0.0004000256

Design A-3 Setup:

	Values to be used in 2^2 Design	
Run	Node 524	Node 1141
1	0.0000496233	0.0003500450
2	0.0000496250	0.0003500450
3	0.0000496233	0.0004000256
4	0.0000496250	0.0004000256
Center	0.0000496242	0.0003750353

SSE Values obtained from previous 2x2 run.

	Coded Values		**Design A-3**
Run	Node 524	Node 1141	SSE
1	-1	-1	0.0001791029
2	1	-1	0.0001791032
3	-1	1	0.0001839788
4	1	1	0.0001839791
Center	0	0	0.0001799550

Calculated Sum of Squares Solution obtained from SSE runs yields the reduction gradient (difference between run #1 and #2) and successive model run parameters.

Run Number	Low Zone Val	High Zone Val	SSE
1	0.0000496242	0.0003750353	0.000179955011
2	0.0000496242	0.0003500450	0.000179103077
3	0.0000496241	0.0003250547	0.000195617125
4	0.0000496241	0.0003000644	0.000186930371
5	0.0000496240	0.0002750741	0.000195617125

Design A Experimental Values for Experiment One
Calculating Flow Rates

Design A-4 Experimental Values:

Node Number	Low Flow Rate	Med Flow Rate	High Flow Rate
524	0.0000496242	0.0000496242	0.0000496242
1141	0.0003500450	0.0003625402	0.0003750353

Design A-4 Setup:

	Values to be used in 2^2 Design	
Run	Node 524	Node 1141
1	0.0000496242	0.0003500450
2	0.0000496242	0.0003500450
3	0.0000496242	0.0003750353
4	0.0000496242	0.0003750353
Center	0.0000496242	0.0003625402

SSE Values obtained from previous 2x2 run.

	Coded Values		**Design A-4**
Run	Node 524	Node 1141	SSE
1	-1	-1	0.000179103077
2	1	-1	0.000179103077
3	-1	1	0.000179955011
4	1	1	0.000179955011
Center	0	0	0.000179132305

Calculated Sum of Squares Solution obtained from SSE runs yields the reduction gradient (difference between run #1 and #2) and successive model run parameters.

Run Number	Low Zone Val	High Zone Val	SSE
1	0.0000496242	0.0003625402	0.000179132305
2	0.0000496242	0.0003500450	0.000179103077
3	0.0000496242	0.0003375498	0.000179867807
4	0.0000496242	0.0003250546	0.000181426964
5	0.0000496242	0.0003125594	0.000183781020

Design A Experimental Values for Experiment One
Calculating Flow Rates

Design A-5 Experimental Values:

Node Number	Low Flow Rate	Med Flow Rate	High Flow Rate
524	0.0000496242	0.0000496242	0.0000496242
1141	0.0003500450	0.0003562926	0.0003625402

Design A-5 Setup:

	Values to be used in 2^2 Design	
Run	Node 524	Node 1141
1	0.0000496242	0.0003500450
2	0.0000496242	0.0003500450
3	0.0000496242	0.0003625402
4	0.0000496242	0.0003625402
Center	0.0000496242	0.0003562926

SSE Values obtained from previous 2x2 run.

	Coded Values		**Design A-5**
Run	Node 524	Node 1141	SSE
1	-1	-1	0.000179103077
2	1	-1	0.000179103077
3	-1	1	0.000179132305
4	1	1	0.000179132305
Center	0	0	0.000179018476

Calculated Sum of Squares Solution obtained from SSE runs yields the reduction gradient (difference between run #1 and #2) and successive model run parameters.

Run Number	Low Zone Val	High Zone Val	SSE
1	0.0000496242	0.0003500450	0.000179103077
2	0.0000496242	0.0003510863	0.000179075196
3	0.0000496242	0.0003521275	0.000179052826
4	0.0000496242	0.0003531688	0.000179035969
5	0.0000496242	0.0003542101	0.000179024626
6	0.0000496242	0.0003552513	0.000179018795
7	0.0000496242	0.0003573339	0.000179023669
8	0.0000496242	0.0003594164	0.000179050590
9	0.0000496242	0.0003614989	0.000179099555
10	0.0000496242	0.0003635815	0.000179170566
11	0.0000496242	0.0003656640	0.000179263613
12	0.0000496242	0.0003677465	0.000179378697

Solution	
Lower Zone	**Upper Zone**
0.000049624	**0.000355**

MATHCAD Worksheets
Experiment #1 RSM RUNS - FIRST ITERATION

Coded Conditions SSE Response Uncoded Conditions

$$A := \begin{bmatrix} -1 & -1 \\ 1 & -1 \\ -1 & 1 \\ 1 & 1 \\ 0 & 0 \end{bmatrix} \quad b := \begin{bmatrix} 0.0001825507 \\ 0.0001858608 \\ 0.000328244 \\ 0.0003322036 \\ 0.0002317981 \end{bmatrix} \quad UnCoded := \begin{bmatrix} .00004 & .0004 \\ .00006 & .0004 \\ .00004 & .0006 \\ .00006 & .0006 \\ .00005 & .0005 \end{bmatrix}$$

Regression solution for X Solution

$$x := \left(A^T \cdot A\right)^{-1} \cdot \left(A^T \cdot b\right) \qquad x = \begin{pmatrix} 1.817425 \cdot 10^{-6} \\ 7.300902 \cdot 10^{-5} \end{pmatrix}$$

Number of Variables - 1

$n := 1$

Normalization Constant

$$N := \sqrt{\sum_{i=0}^{n} \left(x_i\right)^2}$$

Normalized Coded Gradient Equation

$$g := -\frac{x}{N}$$

Solution: Normalized Coded Unit Vectors

$$g = \begin{pmatrix} -0.02488545 \\ -0.99969031 \end{pmatrix}$$

Need to next convert from coded to uncoded values:

$$m1 := \frac{0.00006 - 0.00004}{1 - (-1)} \qquad m2 := \frac{0.0006 - 0.0004}{1 - (-1)}$$

$$m1 = 1 \cdot 10^{-5} \qquad m2 = 1 \cdot 10^{-4}$$

$b1 := .00006 - m1$ $b2 := .0006 - m2$

$b1 = 5 \cdot 10^{-5}$ $b2 = 5 \cdot 10^{-4}$

The FORMULA (First Variable)... converting from coded to uncoded values:

Range $X_1 := m1 \cdot (-1) + b1$ $X_2 := m1 \cdot 0 + b1$ $X_3 := m1 \cdot 1 + b1$

and $X_1 = 4 \cdot 10^{-5}$ $X_2 = 5 \cdot 10^{-5}$ $X_3 = 6 \cdot 10^{-5}$

Center

Conversions check. Formula for 1st variable is:

$f1(x) := m1 \cdot x + b1$

The FORMULA (Second Variable)... converting from coded to uncoded values:

Range $X_4 := m2 \cdot (-1) + b2$ $X_5 := m2 \cdot 0 + b2$ $X_6 := m2 \cdot 1 + b2$

and $X_4 = 4 \bullet 10^{-4}$ $X_5 = 5 \cdot 10^{-4}$ $X_6 = 6 \cdot 10^{-4}$

Center

Conversion checks. Formula for 2nd variable is:

$f2(x) := m2 \cdot x + b2$

These are the converted values for use in the new data runs. I will use the center value MINUS these values until the SSE starts to rise.

$G1 := f1(g_0)$ $G2 := f2(g_1)$

$G1 = 0.00004975$ $G2 = 0.00040003$

$Diff1 := .00005 - G1$ $Diff1 = 0.00000025$

$Diff2 := .0005 - G2$ $Diff2 \quad 0.00009997$

RUN VALUES TO BE USED:

$$RunData := \begin{bmatrix} 0.00005 & 0.0005 \\ 0.00005 - Diff1 & 0.0005 - Diff2 \\ 0.00005 - 2 \cdot Diff1 & 0.0005 - 2 \cdot Diff2 \\ 0.00005 - 3 \cdot Diff1 & 0.0005 - 3 \cdot Diff2 \\ 0.00005 - 4 \cdot Diff1 & 0.0005 - 4 \cdot Diff2 \\ 0.00005 - 5 \cdot Diff1 & 0.0005 - 5 \cdot Diff2 \end{bmatrix}$$

$$RunData = \begin{bmatrix} 0.00005 & 0.0005 \\ 0.00004975 & 0.00040003 \\ 0.0000495 & 0.00030006 \\ 0.00004925 & 0.00020009 \\ 0.000049 & 0.00010012 \\ 0.00004876 & 0.00000015 \end{bmatrix}$$

Experiment #1 RSM RUNS - SECOND ITERATION

Coded Conditions

$$A := \begin{bmatrix} -1 & -1 \\ 1 & -1 \\ -1 & 1 \\ 1 & 1 \\ 0 & 0 \end{bmatrix}$$

SSE Response

$$b := \begin{bmatrix} 0.0001869133 \\ 0.0001869503 \\ 0.0001839596 \\ 0.0001840007 \\ 0.0001791032 \end{bmatrix}$$

Uncoded Conditions

$$UnCoded := \begin{bmatrix} 0.00004950 & 0.00030006 \\ 0.00004975 & 0.00030006 \\ 0.00004950 & 0.00040003 \\ 0.00004975 & 0.00040003 \\ 0.000049625 & 0.000350045 \end{bmatrix}$$

Regression solution for X

$$x := \left(A^T \cdot A\right)^{-1} \cdot \left(A^T \cdot b\right)$$

Solution

$$x = \begin{pmatrix} 1.9525 \bullet 10^{-8} \\ -1.47583 \bullet 10^{-6} \end{pmatrix}$$

Number of Variables - 1

$$n := 1$$

Normalization Constant

$$N := \sqrt{\sum_{i=0}^{n} \left(x_i\right)^2}$$

Normalized Coded Gradient Equation

$$g := -\frac{x}{N}$$

Solution: Normalized Coded Unit Vectors

$$g = \begin{pmatrix} -0.01322873 \\ 0.9999125 \end{pmatrix}$$

Conversion from coded to uncoded values:

$$m1 := \frac{0.00004975 - 0.00004950}{1 - (-1)} \qquad m2 := \frac{0.00040003 - 0.00030006}{1 - (-1)}$$

$$m1 = 1.25 \cdot 10^{-7} \text{ ■} \qquad m2 = 4.9985 \cdot 10^{-5} \text{ ■}$$

$$b1 := 0.00004975 - m1 \qquad b2 := 0.00040003 - m2$$

$$b1 = 4.9625 \cdot 10^{-5} \text{ ■} \qquad b2 = 3.50045 \cdot 10^{-4} \text{ ■}$$

The FORMULA (First Variable)... converting from coded to uncoded values:

Range $X_1 := m1 \cdot (-1) + b1$ $X_2 := m1 \cdot 0 + b1$ $X_3 := m1 \cdot 1 + b1$

and $X_1 = 4.9500000000\,10^{-5}$ ■ $X_2 = 4.9625 \cdot 10^{-5}$ ■ $X_3 = 4.975 \cdot 10^{-5}$ ■

Center

Conversion checks. Formula for 1st variable is:

$$f1(x) := m1 \cdot x + b1$$

The FORMULA (Second Variable)... converting from coded to uncoded values:

Range $X_4 := m2 \cdot (-1) + b2$ $X_5 := m2 \cdot 0 + b2$ $X_6 := m2 \cdot 1 + b2$

and Center $X_4 = 3.0006000000 \cdot 10^{-4}$ ▮ $X_5 = 3.50045 \cdot 10^{-4}$ ▮ $X_6 = 4.0003 \cdot 10^{-4}$ ▮

Conversion checks. Formula for 2nd variable is:

$f2(x) := m2 \cdot x + b2$

These are the converted values for use in the new data runs. I will use the center value MINUS these values until the SSE starts to rise.

$G1 := f1(g_0)$ $G2 := f2(g_1)$

$G1 = 0.00004962$ ▮ $G2 = 0.00040003$ ▮

$Diff1 := 0.000049625 - G1$ $Diff1 = 1.65359134 \cdot 10^{-9}$ ▮

$Diff2 := 0.000350045 - G2$ $Diff2 = -0.00004998$ ▮

RUN VALUES TO BE USED:

$$RunData := \begin{bmatrix} 0.000049625 & 0.000350045 \\ 0.000049625 - Diff1 & 0.000350045 - Diff2 \\ 0.000049625 - 2 \cdot Diff1 & 0.000350045 - 2 \cdot Diff2 \\ 0.000049625 - 3 \cdot Diff1 & 0.000350045 - 3 \cdot Diff2 \\ 0.000049625 - 4 \cdot Diff1 & 0.000350045 - 4 \cdot Diff2 \end{bmatrix}$$

$$RunData = \begin{bmatrix} 0.000049625 & 0.000350045 \\ 0.0000496233 & 0.0004000256 \\ 0.0000496217 & 0.0004500063 \\ 0.00004962 & 0.0004999869 \\ 0.0000496184 & 0.0005499675 \end{bmatrix}$$ ▮

Experiment #1 RSM RUNS - THIRD ITERATION

Coded Conditions SSE Response Uncoded Conditions

$$A := \begin{bmatrix} -1 & -1 \\ 1 & -1 \\ -1 & 1 \\ 1 & 1 \\ 0 & 0 \end{bmatrix} \quad b := \begin{bmatrix} 0.000179102937 \\ 0.000179103202 \\ 0.000183978845 \\ 0.000183979124 \\ 0.000179955011 \end{bmatrix} \quad UnCoded := \begin{bmatrix} 0.0000496233 & 0.000350045 \\ 0.000049625 & 0.000350045 \\ 0.0000496233 & 0.0004000256 \\ 0.000049625 & 0.0004000256 \\ 0.0000496242 & 0.0003750353 \end{bmatrix}$$

Regression solution for X

$$x := \left(A^T \cdot A\right)^{-1} \cdot \left(A^T \cdot b\right)$$

Solution

$$x = \begin{pmatrix} 1.36 \cdot 10^{-10} \\ 2.4379575 \cdot 10^{-6} \end{pmatrix}$$

Number of Variables - 1

$$n := 1$$

Normalization Constant

$$N := \sqrt{\sum_{i=0}^{n} \left(x_i\right)^2}$$

Normalized Coded Gradient Equation

$$g := -\frac{x}{N}$$

Solution: Normalized Coded Unit Vectors

$$g = \begin{pmatrix} -5.57844014 \cdot 10^{-5} \\ -1 \end{pmatrix}$$

Need to next convert from coded to uncoded values:

$$m1 := \frac{UnCoded_{1,0} - UnCoded_{0,0}}{1 - (-1)} \qquad m2 := \frac{UnCoded_{2,1} - UnCoded_{0,1}}{1 - (-1)}$$

$$m1 = 8.5 \cdot 10^{-10} \qquad m2 = 2.49903 \cdot 10^{-5}$$

$$b1 := UnCoded_{1,0} - m1 \qquad b2 := UnCoded_{2,1} - m2$$

$$b1 = 4.9 \qquad b2 = 3.7$$

The FORMULA (First Variable)... converting from coded to uncoded values:

Range and Center

$$X_1 := m1 \cdot (-1) + b1 \qquad X_2 := m1 \cdot 0 + b1 \qquad X_3 := m1 \cdot 1 + b1$$

$$X_1 = 4.9623300000 \cdot 10^{-5} \qquad X_2 = 4.962415 \cdot 10^{-5} \qquad X_3 = 4.9625 \cdot 10^{-5}$$

Conversion checks. Formula for 1st variable is:

$$f1(x) := m1 \cdot x + b1$$

The FORMULA (Second Variable)... converting from coded to uncoded values:

Range $X_4 := m2 \cdot (-1) + b2$ $X_5 := m2 \cdot 0 + b2$ $X_6 := m2 \cdot 1 + b2$

and Center $X_4 = 3.5004500000 \cdot 10^{-4}$ ▮ $X_5 = 3.750353 \cdot 10^{-4}$ ▮ $X_6 = 4.000256 \cdot 10^{-4}$ ▮

Conversion checks. Formula for 2nd variable is:

$f2(x) := m2 \cdot x + b2$

These are the converted values for use in the new data runs. I will use the center value MINUS these values until the SSE starts to rise.

$G1 := f1(g_0)$ $G2 := f2(g_1)$

$G1 = 0.00004962$ ▮ $G2 = 0.00035005$ ▮

$Diff1 := UnCoded_{4,0} - G1$ $Diff2 := UnCoded_{4,1} - G2$

$Diff1 = 5.00474167 \cdot 10^{-11}$ ▮ $Diff2 = 0.00002499$ ▮

RUN VALUES TO BE USED:

$$RunData := \begin{bmatrix} UnCoded_{4,0} & UnCoded_{4,1} \\ UnCoded_{4,0} - Diff1 & UnCoded_{4,1} - Diff2 \\ UnCoded_{4,0} - 2 \cdot Diff1 & UnCoded_{4,1} - 2 \cdot Diff2 \\ UnCoded_{4,0} - 3 \cdot Diff1 & UnCoded_{4,1} - 3 \cdot Diff2 \\ UnCoded_{4,0} - 4 \cdot Diff1 & UnCoded_{4,1} - 4 \cdot Diff2 \end{bmatrix}$$

$$RunData = \begin{bmatrix} 0.0000496242 & 0.0003750353 \\ 0.00004962415 & 0.000350045 \\ 0.0000496241 & 0.0003250547 \\ 0.00004962405 & 0.0003000644 \\ 0.000049624 & 0.0002750741 \end{bmatrix}$$ ▮

Experiment #1 RSM RUNS - FOURTH ITERATION

Coded Conditions

$$A := \begin{bmatrix} -1 & -1 \\ 1 & -1 \\ -1 & 1 \\ 1 & 1 \\ 0 & 0 \end{bmatrix}$$

SSE Response

$$b := \begin{bmatrix} 0.000179103077 \\ 0.000179103077 \\ 0.000179955011 \\ 0.000179955011 \\ 0.000179132305 \end{bmatrix}$$

Uncoded Conditions

$$UnCoded := \begin{bmatrix} 0.0000496242 & 0.000350045 \\ 0.0000496242 & 0.000350045 \\ 0.0000496242 & 0.0003750353 \\ 0.0000496242 & 0.0003750353 \\ 0.0000496242 & 0.0003625402 \end{bmatrix}$$

Regression solution for X

$$x := \left(A^T \cdot A\right)^{-1} \cdot \left(A^T \cdot b\right)$$

Solution

$$x = \begin{pmatrix} 0 \\ 4.25967 \cdot 10^{-7} \end{pmatrix}$$

Number of Variables - 1

$$n := 1$$

Normalization Constant

$$N := \sqrt{\sum_{i=0}^{n} \left(x_i\right)^2}$$

Normalized Coded Gradient Equation

$$g := -\frac{x}{N}$$

Solution: Normalized Coded Unit Vectors

$$g = \begin{pmatrix} 0 \\ -1 \end{pmatrix}$$

Need to next convert from coded to uncoded values:

$$m1 := \frac{UnCoded_{1,0} - UnCoded_{0,0}}{1 - (-1)} \qquad m2 := \frac{UnCoded_{2,1} - UnCoded_{0,1}}{1 - (-1)}$$

$$m1 = 0 \qquad m2 = 1.249515 \cdot 10^{-5}$$

$$b1 := UnCoded_{1,0} - m1 \qquad b2 := UnCoded_{2,1} - m2$$

$$b1 = 4.96242 \cdot 10^{-5} \qquad b2 = 3.6254015 \cdot 10^{-4}$$

The FORMULA (First Variable)... converting from coded to uncoded values:

Range and Center

$$X_1 := m1 \cdot (-1) + b1 \qquad X_2 := m1 \cdot 0 + b1 \qquad X_3 := m1 \cdot 1 + b1$$

$$X_1 = 4.9624200000 \cdot 10^{-5} \qquad X_2 = 4.96242 \cdot 10^{-5} \qquad X_3 = 4.96242 \cdot 10^{-5}$$

Conversion checks. Formula for 1st variable is:

$$f1(x) := m1 \cdot x + b1$$

The FORMULA (Second Variable)... converting from coded to uncoded values:

Range $X_4 := m2 \cdot (-1) + b2$ $X_5 := m2 \cdot 0 + b2$ $X_6 := m2 \cdot 1 + b2$

and $X_4 = 3.5004500000 \cdot 10^{-4}$ $X_5 = 3.6254015 \cdot 10^{-4}$ $X_6 = 3.750353 \cdot 10^{-4}$

Center

Conversion checks. Formula for 2nd variable is:

$f2(x) := m2 \cdot x + b2$

These are the converted values for use in the new data runs. I will use the center value MINUS these values until the SSE starts to rise.

$G1 := f1(g_0)$ $G2 := f2(g_1)$

$G1 = 0.00004962$ $G2 = 0.00035005$

$Diff1 := UnCoded_{4,0} - G1$ $Diff2 := UnCoded_{4,1} - G2$

$Diff1 = 0$ $Diff2 = 0.0000125$

RUN VALUES TO BE USED:

$$RunData := \begin{bmatrix} UnCoded_{4,0} & UnCoded_{4,1} \\ UnCoded_{4,0} - Diff1 & UnCoded_{4,1} - Diff2 \\ UnCoded_{4,0} - 2 \cdot Diff1 & UnCoded_{4,1} - 2 \cdot Diff2 \\ UnCoded_{4,0} - 3 \cdot Diff1 & UnCoded_{4,1} - 3 \cdot Diff2 \\ UnCoded_{4,0} - 4 \cdot Diff1 & UnCoded_{4,1} - 4 \cdot Diff2 \end{bmatrix}$$

$$RunData = \begin{bmatrix} 0.0000496242 & 0.0003625402 \\ 0.0000496242 & 0.000350045 \\ 0.0000496242 & 0.0003375498 \\ 0.0000496242 & 0.0003250546 \\ 0.0000496242 & 0.0003125594 \end{bmatrix}$$

Experiment #1 RSM RUNS - FIFTH ITERATION

Coded Conditions

$$A := \begin{bmatrix} -1 & -1 \\ 1 & -1 \\ -1 & 1 \\ 1 & 1 \\ 0 & 0 \end{bmatrix}$$

SSE Response

$$b := \begin{bmatrix} 0.000179103077 \\ 0.000179103077 \\ 0.000179132305 \\ 0.000179132305 \\ 0.000179018476 \end{bmatrix}$$

Uncoded Conditions

$$UnCoded := \begin{bmatrix} 0.0000496242 & 0.0003500450 \\ 0.0000496242 & 0.0003500450 \\ 0.0000496242 & 0.0003625402 \\ 0.0000496242 & 0.0003625402 \\ 0.0000496242 & 0.0003562926 \end{bmatrix}$$

Regression solution for X

$$x := \left(A^T \cdot A\right)^{-1} \cdot \left(A^T \cdot b\right)$$

Solution

$$x = \begin{pmatrix} 0 \\ 1.4614 \cdot 10^{-8} \end{pmatrix}$$

Number of Variables - 1

$$n := 1$$

Normalization Constant

$$N := \sqrt{\sum_{i=0}^{n} \left(x_i\right)^2}$$

Normalized Coded Gradient Equation

$$g := -\frac{x}{N}$$

Solution: Normalized Coded Unit Vectors

$$g = \begin{pmatrix} 0 \\ -1 \end{pmatrix}$$

Need to next convert from coded to uncoded values:

$$m1 := \frac{UnCoded_{1,0} - UnCoded_{0,0}}{1 - (-1)} \qquad m2 := \frac{UnCoded_{2,1} - UnCoded_{0,1}}{1 - (-1)}$$

$$m1 = 0 \qquad m2 = 6.2476 \cdot 10^{-6}$$

$$b1 := UnCoded_{1,0} - m1 \qquad b2 := UnCoded_{2,1} - m2$$

$$b1 = 4.96242 \cdot 10^{-5} \qquad b2 = 3.562926 \cdot 10^{-4}$$

The FORMULA (First Variable)... converting from coded to uncoded values:

Range $X_1 := m1 \cdot (-1) + b1$ $X_2 := m1 \cdot 0 + b1$ $X_3 := m1 \cdot 1 + b1$

and $X_1 = 4.9624200000 \cdot 10^{-5}$ $X_2 = 4.96242 \cdot 10^{-5}$ $X_3 = 4.96242 \cdot 10^{-5}$

Center

Conversion checks. Formula for 1st variable is:

$$f1(x) := m1 \cdot x + b1$$

The FORMULA (Second Variable)... converting from coded to uncoded values:

Range $X_4 := m2 \cdot (-1) + b2$ $X_5 := m2 \cdot 0 + b2$ $X_6 := m2 \cdot 1 + b2$

and $X_4 = 3.5004500000 \cdot 10^{-4}$, $X_5 = 3.562926 \cdot 10^{-4}$, $X_6 = 3.625402 \cdot 10^{-4}$,

Center

Conversion checks. Formula for 2nd variable is:

$f2(x) := m2 \cdot x + b2$

These are the converted values for use in the new data runs. I will use the center value MINUS these values until the SSE starts to rise.

$G1 := f1(g_0)$ $G2 := f2(g_1)$

$G1 = 0.00004962$, $G2 = 0.00035005$,

$Diff1 := UnCoded_{4,0} - G1$ $Diff2 := UnCoded_{4,1} - G2$

$Diff1 = 0$ $Diff2 = 0.00000625$,

RUN VALUES TO BE USED:

$$RunData := \begin{bmatrix} UnCoded_{4,0} & UnCoded_{4,1} \\ UnCoded_{4,0} - Diff1 & UnCoded_{4,1} - Diff2 \\ UnCoded_{4,0} - 2 \cdot Diff1 & UnCoded_{4,1} - 2 \cdot Diff2 \\ UnCoded_{4,0} - 3 \cdot Diff1 & UnCoded_{4,1} - 3 \cdot Diff2 \\ UnCoded_{4,0} - 4 \cdot Diff1 & UnCoded_{4,1} - 4 \cdot Diff2 \end{bmatrix}$$

$$RunData = \begin{bmatrix} 0.0000496242 & 0.0003562926 \\ 0.0000496242 & 0.000350045 \\ 0.0000496242 & 0.0003437974 \\ 0.0000496242 & 0.0003375498 \\ 0.0000496242 & 0.0003313022 \end{bmatrix} ,$$

Graphical Representation of Experiment #1 SSE Runs

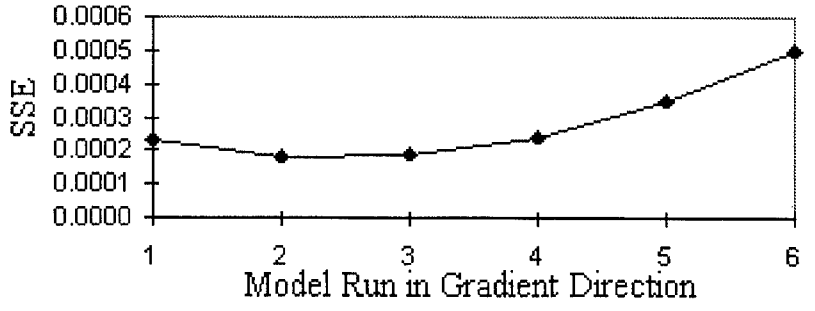


Figure B-1. Experiment #1 calculation of first iteration of SSE by run number.

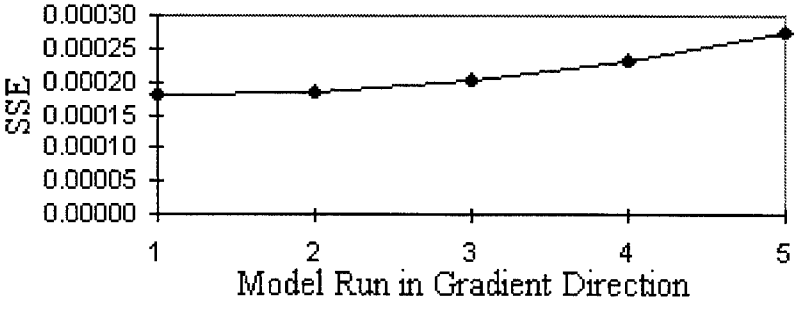


Figure B-2. Experiment #1 calculation of second iteration of SSE by run number.

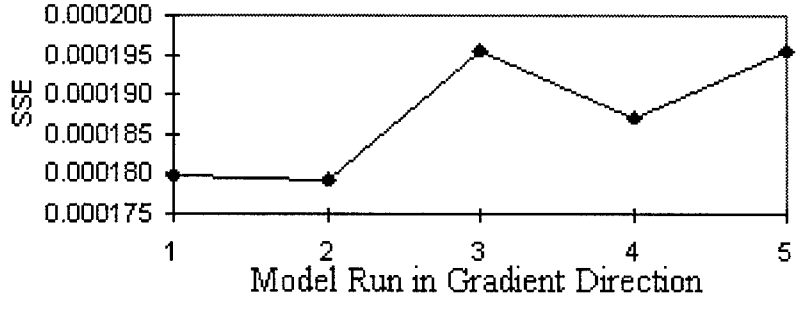


Figure B-3. Experiment #1 calculation of third iteration of SSE by run number.

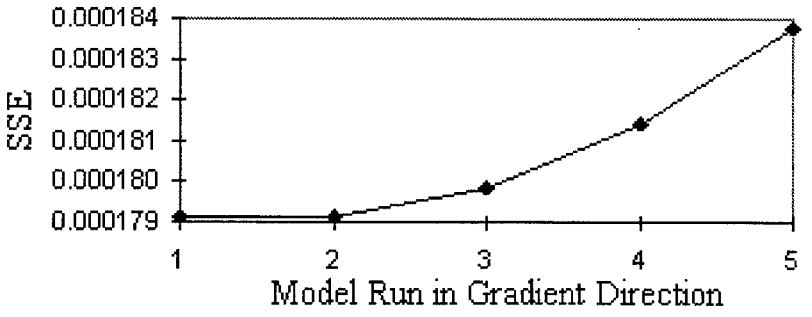


Figure B-4. Experiment #1 calculation of fourth iteration of SSE by run number.

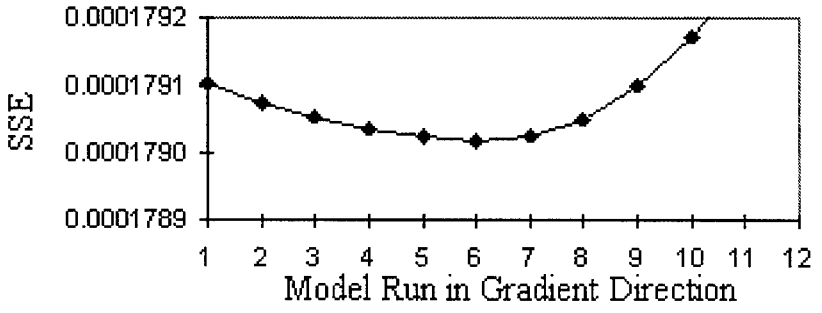


Figure B-5. Experiment #1 calculation of fifth iteration of SSE by run number.

SAMPLE SUTRA INPUT FILE

```
SUTRA SOLUTE TRANSPORT   HILL AIR FORCE BASE MODEL  ..SCHEMATIC..
***** HILL AIR FORCE BASE GW/CT...SCHEMATIC
REPRESENTATION...SINGLE 1000 yr S
++ EXAMPLE RUN FOR SUTRA DOCUMENTATION - SECTION 6.4, PAGE 188 ++
 1681 1600   85    0   82   00    1    0   00   99
    0   +1   +0   +1  999
   0.00000       1.00D-00       1.00D-00
  300       1.7882+06       5.3646+08    999999       1.0
3.15577+09    1   01
   10    0    0    0   00   00   +1    1
         01
  0.000E-0    0001.0       0.0    1.0E00      0.00   0.00   1.000
     0.0E0     0.0E0     0.0E0    0001.0
NONE
    0.0E-9    0.0E-9    0.0E-9    0.0E-9
      00.0      00.0
NODE            30.48     30.48     10.00      0.20
    1          0.0000    0.0000    1.000    1.0000
    2          1.0000    0.0000    1.000    1.0000
    3          2.0000    0.0000    1.000    1.0000
    4          3.0000    0.0000    1.000    1.0000
    5          4.0000    0.0000    1.000    1.0000
    6          5.0000    0.0000    1.000    1.0000
    7          6.0000    0.0000    1.000    1.0000
    8          7.0000    0.0000    1.000    1.0000
    9          8.0000    0.0000    1.000    1.0000
   10          9.0000    0.0000    1.000    1.0000
    .
    .
    .
 1670         29.0000   40.0000    1.000    1.0000
 1671         30.0000   40.0000    1.000    1.0000
 1672         31.0000   40.0000    1.000    1.0000
 1673         32.0000   40.0000    1.000    1.0000
 1674         33.0000   40.0000    1.000    1.0000
 1675         34.0000   40.0000    1.000    1.0000
 1676         35.0000   40.0000    1.000    1.0000
 1677         36.0000   40.0000    1.000    1.0000
 1678         37.0000   40.0000    1.000    1.0000
 1679         38.0000   40.0000    1.000    1.0000
 1680         39.0000   40.0000    1.000    1.0000
 1681         40.0000   40.0000    1.000    1.0000
```

ELEMENT	9.445E-8	9.445E-8	0.000	100.0	100.0	30.00	30.0
1	1.00D+00	1.00D+00	1.00D+00	1.000	1.000	1.0000	1.0000
2	1.00D+00	1.00D+00	1.00D+00	1.000	1.000	1.0000	1.0000
3	1.00D+00	1.00D+00	1.00D+00	1.000	1.000	1.0000	1.0000
4	1.00D+00	1.00D+00	1.00D+00	1.000	1.000	1.0000	1.0000
5	1.00D+00	1.00D+00	1.00D+00	1.000	1.000	1.0000	1.0000
6	1.00D+00	1.00D+00	1.00D+00	1.000	1.000	1.0000	1.0000
7	1.00D+00	1.00D+00	1.00D+00	1.000	1.000	1.0000	1.0000
8	1.00D+00	1.00D+00	1.00D+00	1.000	1.000	1.0000	1.0000
9	1.00D+00	1.00D+00	1.00D+00	1.000	1.000	1.0000	1.0000
10	1.00D+00	1.00D+00	1.00D+00	1.000	1.000	1.0000	1.0000
11	0.50D+00	0.50D+00	1.00D+00	1.000	1.000	1.0000	1.0000
12	0.50D+00	0.50D+00	1.00D+00	1.000	1.000	1.0000	1.0000
13	0.50D+00	0.50D+00	1.00D+00	1.000	1.000	1.0000	1.0000
14	0.50D+00	0.50D+00	1.00D+00	1.000	1.000	1.0000	1.0000
15	0.50D+00	0.50D+00	1.00D+00	1.000	1.000	1.0000	1.0000
.							
.							
.							
1581	1.30D+01	1.30D+01	1.00D+00	1.000	1.000	1.0000	1.0000
1582	1.30D+01	1.30D+01	1.00D+00	1.000	1.000	1.0000	1.0000
1583	1.30D+01	1.30D+01	1.00D+00	1.000	1.000	1.0000	1.0000
1584	1.30D+01	1.30D+01	1.00D+00	1.000	1.000	1.0000	1.0000
1585	1.00D+00	1.00D+00	1.00D+00	1.000	1.000	1.0000	1.0000
1586	1.00D+00	1.00D+00	1.00D+00	1.000	1.000	1.0000	1.0000
1587	1.00D+00	1.00D+00	1.00D+00	1.000	1.000	1.0000	1.0000
1588	1.00D+00	1.00D+00	1.00D+00	1.000	1.000	1.0000	1.0000
1589	1.00D+00	1.00D+00	1.00D+00	1.000	1.000	1.0000	1.0000
1590	1.00D+00	1.00D+00	1.00D+00	1.000	1.000	1.0000	1.0000
1591	1.00D+00	1.00D+00	1.00D+00	1.000	1.000	1.0000	1.0000
1592	1.00D+00	1.00D+00	1.00D+00	1.000	1.000	1.0000	1.0000
1593	1.00D+00	1.00D+00	1.00D+00	1.000	1.000	1.0000	1.0000
1594	0.10D+00	0.10D+00	1.00D+00	1.000	1.000	1.0000	1.0000
1595	0.10D+00	0.10D+00	1.00D+00	1.000	1.000	1.0000	1.0000
1596	0.10D+00	0.10D+00	1.00D+00	1.000	1.000	1.0000	1.0000
1597	0.10D+00	0.10D+00	1.00D+00	1.000	1.000	1.0000	1.0000
1598	0.10D+00	0.10D+00	1.00D+00	1.000	1.000	1.0000	1.0000
1599	0.10D+00	0.10D+00	1.00D+00	1.000	1.000	1.0000	1.0000
1600	1.00D+00	1.00D+00	1.00D+00	1.000	1.000	1.0000	1.0000

```
      1133   +10.00000E-4        1.0E-00
0000000000000000000000000000000000000000000000000000000000000000000000000000
    1                0.0E-1                00.0E-0
   42                0.0E-1                00.0E-0
   83                0.0E-1                00.0E-0
  124                0.0E-1                00.0E-0
  165                0.0E-1                00.0E-0
  206                0.0E-1                00.0E-0
  247                0.0E-1                00.0E-0
  288                0.0E-1                00.0E-0
  329                0.0E-1                00.0E-0
  370                0.0E-1                00.0E-0
  411                0.0E-1                00.0E-0
  452                0.0E-1                00.0E-0
  493                0.0E-1                00.0E-0
  534                0.0E-1                00.0E-0
  575                0.0E-1                00.0E-0
  616                0.0E-1                00.0E-0
  657                0.0E-1                00.0E-0
  698                0.0E-1                00.0E-0
  739                0.0E-1                00.0E-0
  780                0.0E-1                00.0E-0
  821                0.0E-1                00.0E-0
  862                0.0E-1                00.0E-0
  903                0.0E-1                00.0E-0
  944                0.0E-1                00.0E-0
  985                0.0E-1                00.0E-0
 1026                0.0E-1                00.0E-0
 1067                0.0E-1                00.0E-0
 1108                0.0E-1                00.0E-0
 1149                0.0E-1                00.0E-0
 1190                0.0E-1                00.0E-0
 1231                0.0E-1                00.0E-0
 1272                0.0E-1                00.0E-0
 1313                0.0E-1                00.0E-0
 1354                0.0E-1                00.0E-0
 1395                0.0E-1                00.0E-0
 1436                0.0E-1                00.0E-0
 1477                0.0E-1                00.0E-0
 1518                0.0E-1                00.0E-0
 1559                0.0E-1                00.0E-0
 1600                0.0E-1                00.0E-0
 1641                0.0E-1                00.0E-0
   41                3.0E+5                00.0E-3
   82                3.0E+5                00.0E-3
  123                3.0E+5                00.0E-3
  164                3.0E+5                00.0E-3
  205                3.0E+5                00.0E-3
  246                3.0E+5                00.0E-3
  287                3.0E+5                00.0E-3
  328                3.0E+5                00.0E-3
  369                3.0E+5                00.0E-3
  410                3.0E+5                00.0E-3
  451                3.0E+5                00.0E-3
  492                3.0E+5                00.0E-3
  533                3.0E+5                00.0E-3
```

```
  574                3.0E+5                00.0E-3
  615                3.0E+5                00.0E-3
  656                3.0E+5                00.0E-3
  697                3.0E+5                00.0E-3
  738                3.0E+5                00.0E-3
  779                3.0E+5                00.0E-3
  820                3.0E+5                00.0E-3
  861                3.0E+5                00.0E-3
  902                3.0E+5                00.0E-3
  943                3.0E+5                00.0E-3
  984                3.0E+5                00.0E-3
 1025                3.0E+5                00.0E-3
 1066                3.0E+5                00.0E-3
 1107                3.0E+5                00.0E-3
 1148                3.0E+5                00.0E-3
 1189                3.0E+5                00.0E-3
 1230                3.0E+5                00.0E-3
 1271                3.0E+5                00.0E-3
 1312                3.0E+5                00.0E-3
 1353                3.0E+5                00.0E-3
 1394                3.0E+5                00.0E-3
 1435                3.0E+5                00.0E-3
 1476                3.0E+5                00.0E-3
 1517                3.0E+5                00.0E-3
 1558                3.0E+5                00.0E-3
 1599                3.0E+5                00.0E-3
 1640                3.0E+5                00.0E-3
 1681                3.0E+5                00.0E-3
0000000000000000000000000000000000000000000000000000000000000000000000000000
    1     1     2    43    42
    2     2     3    44    43
    3     3     4    45    44
    4     4     5    46    45
    5     5     6    47    46
      .
      .
      .
 1586  1625  1626  1667  1666
 1587  1626  1627  1668  1667
 1588  1627  1628  1669  1668
 1589  1628  1629  1670  1669
 1590  1629  1630  1671  1670
 1591  1630  1631  1672  1671
 1592  1631  1632  1673  1672
 1593  1632  1633  1674  1673
 1594  1633  1634  1675  1674
 1595  1634  1635  1676  1675
 1596  1635  1636  1677  1676
 1597  1636  1637  1678  1677
 1598  1637  1638  1679  1678
 1599  1638  1639  1680  1679
 1600  1639  1640  1681  1680
```

PLACKETT-BURMAN DESIGN FOR 19 NODES

Node Number

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
Run 1	1	1	-1	-1	1	1	1	1	-1	1	-1	1	-1	-1	-1	-1	1	1	-1
Run 2	-1	1	1	-1	-1	1	1	1	1	-1	1	-1	1	-1	-1	-1	-1	1	1
Run 3	1	-1	1	1	-1	-1	1	1	1	1	-1	1	-1	1	-1	-1	-1	-1	1
Run 4	1	1	-1	1	1	-1	-1	1	1	1	1	-1	1	-1	1	-1	-1	-1	-1
Run 5	-1	1	1	-1	1	1	-1	-1	1	1	1	1	-1	1	-1	1	-1	-1	-1
Run 6	-1	-1	1	1	-1	1	1	-1	-1	1	1	1	1	-1	1	-1	1	-1	-1
Run 7	-1	-1	-1	1	1	-1	1	1	-1	-1	1	1	1	1	-1	1	-1	1	-1
Run 8	-1	-1	-1	-1	1	1	-1	1	1	-1	-1	1	1	1	1	-1	1	-1	1
Run 9	1	-1	-1	-1	-1	1	1	-1	1	1	-1	-1	1	1	1	1	-1	1	-1
Run 10	-1	1	-1	-1	-1	-1	1	1	-1	1	1	-1	-1	1	1	1	1	-1	1
Run 11	1	-1	1	-1	-1	-1	-1	1	1	-1	1	1	-1	-1	1	1	1	1	-1
Run 12	-1	1	-1	1	-1	-1	-1	-1	1	1	-1	1	1	-1	-1	1	1	1	1
Run 13	1	-1	1	-1	1	-1	-1	-1	-1	1	1	-1	1	1	-1	-1	1	1	1
Run 14	1	1	-1	1	-1	1	-1	-1	-1	-1	1	1	-1	1	1	-1	-1	1	1
Run 15	1	1	1	-1	1	-1	1	-1	-1	-1	-1	1	1	-1	1	1	-1	-1	1
Run 16	1	1	1	1	-1	1	-1	1	-1	-1	-1	-1	1	1	-1	1	1	-1	-1
Run 17	-1	1	1	1	1	-1	1	-1	1	-1	-1	-1	-1	1	1	-1	1	1	-1
Run 18	-1	-1	1	1	1	1	-1	1	-1	1	-1	-1	-1	-1	1	1	-1	1	1
Run 19	1	-1	-1	1	1	1	1	-1	1	-1	1	-1	-1	-1	-1	1	1	-1	1
Run 20	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1

Appendix C. Additional Reading

The following other research assisted me with both environmental engineering basics and the practical use of fuzzy logic.

Dahab, Lee and Bogardi's papers on *A rule-based fuzzy set approach to risk analysis of nitrate-contaminated groundwater* gave insight into how defuzzification techniques--techniques for obtaining real solutions from fuzzy solution spaces--might be replaced by a *rule-based* system to determine risk (Dahab, 1994).

Slichter's article in Ground Water discussed the general laws of the flow of ground waters and was an aid to understanding Darcy's Law (Slichter, 1995).

Baffaut and Chameau's paper on the *Estimation of pollutant loads with fuzzy sets* outlined the uncertainty in estimating pollutant loads and concentrations in runoff using fuzzy sets (Baffaut and Chameau, 1990).

Moon's paper *On mathematical representation and integration of multiple spatial geoscience data sets* described and made comparisons of the three approaches to representing geographical information, the basic probabilistic approach, evidential belief approach (Dempster-Shafer method) and the fuzzy logic approach (Moon, 1993).

Cox's prelude to The Fuzzy Handbook entitled *Solving Problems with Fuzzy Logic* provided software and additional rules for modeling and solving fuzzy logic problems (Cox, 1992).

Bibliography

American Society for Testing & Materials (ASTM), *1996 Annual Book of ASTM Standards,* ASTM, Philadelphia, 1996.

Baffaut, C. and Chameau, J.L., *Estimation of pollutant loads with fuzzy sets*, Civil Engineering Systems, vol. 7(1), pp. 51-61, 1990.

Bedient P.B., Rifai, H.S., and Newell, C.J., *Groundwater Contamination Transport and Remediation*, Prentice-Hall, Inc., Englewood Cliffs, NJ, 1994.

Box, G.E. and Draper, N.R., *Empirical Model Building and Response Surfaces*, New York, NY, Wiley and Sons, 1987.

Carrera, J., State of the art of the inverse problem applied to the flow and solute transport equations, in *Groundwater Flow and Quality Modeling, Proc. NATO Advanced Research Workshop on Advances in Analytical and Numerical Groundwater Flow and Quality Modeling,* Reidel Publishing, Boston, MA, 1987.

Cohen, R. and Mercer, J.W., *DNAPL Site Evaluation*, Boca Raton, Florida, CRC Press, Inc., 1993.

Cornell, J.A., *How to Apply Response Surface Methodology*, Milwaukee, WI, American Society for Quality Control, 1990.

Cotman, R.M., *Groundwater Model Parameter Estimation Using Response Surface Methodology*, Masters Thesis, Air Force Institute of Technology, 1995.

Cox, E., *Solving problems with fuzzy logic*, AI Expert, vol. 7(3), pp. 28-37, 1992.

Dahab, M., Lee, Y., and Bogardi, Y., *A rule-based fuzzy set approach to risk analysis of nitrate-contaminated groundwater*, Water Science Technology, vol. 30(7), pp. 45-52, 1994.

Domenico, P.A. and Schwartz, F.W., *Physical and Chemical Hydrogeology,* John Wiley and Sons, Inc., New York NY, Chapter 19.6, 1990.

EPA, Plume definition, in *Guidance on remedial actions for contaminated groundwater at Superfund sites*, Washington, DC, Office of Emergency and Remedial Response, U.S. Environmental Protection Agency, sec. 3.3.2.2, pg. 3-13, 1989.

Ford, P.J. and Turina, P.J., *Characterization of Hazardous Waste Sites: A Methods Manual, (Site Investigations)*, U.S. Environmental Protection Agency, EPA-600/4-84/075, vol. 1, chapters 4,5,7, April 1985.

Franke, O.L., Reilly, T.E., and Bennett, G.D., *Definition of Boundary and Initial Conditions in the Analysis of Saturated Ground-water Flow Systems -- An Introduction*, U.S. Geological Survey, Techniques of Water Resources Investigations -- Book 3, Government Printing Office, Washington DC, Chapter B5, 1987.

Freeberg, K.M., Bedient, P.B., and Connor, J.A., *Modeling of TCE Contamination and Recovery in a Shallow Sand Acquifer*, Ground Water, vol. 25(1), pp. 70-80, 1987.

Gelhar, L.W., *Stochastic Subsurface Hydrology*, Prentice Hall, Englewood Cliffs, N.J., 1993.

Heiderscheidt, J.L., *Development of Site Characterization Simulator Specifications*, Masters Thesis, Air Force Institute of Technology, Dec 1996.

Hess, K.M., Wolf, S.H., Celia, M.A., *Large-Scale Natural Gradient Tracer Test in Sand and Gravel. 3. Hydraulic Conductivity Variability and Calculated Macrodispersivities*, Water Resources Research, vol. 28(8), pp. 2011-2027, 1992.

Kandel, A., *Fuzzy mathematical techniques with applications*, Addison-Wesley, Reading, MA, 1986.

Khan, I., *Inverse problem in ground water: model development*, Ground Water, vol. 24, pp. 32-48, 1986.

Mackay, D.M., Roberts, P.V., and Cherry, J.A., *Transport of organic contaminants in groundwater*, Environmental Science and Technology, vol. 19(5), pp. 384-392, 1985.

Melville, J.G., Molz, F.J., Given, O., and Widdowson, M.A., *Multilevel Slug Tests with Comparisons to Tracer Data*, Ground Water, vol. 29(6), pp. 897-907, 1991.

Montgomery, D.C., *Introduction to Statistical Quality Control*, John Wiley & Sons, New York, NY, 1991.

Montgomery-Watson, Inc., *Final Phase II Remedial Investigation Report for Operable Unit 3 For Hill Air Force Base, Utah*, March 1995.

Moon, W., *On Mathematical Representation And Integration Of Multiple Spatial Geoscience Data Sets*, Canadian Journal of Remote Sensing, vol. 19(1), pp. 63-67, 1993.

Poulsen, M.M., and Kueper, B.H., *A field experiment to study the behavior of tetrachloroehylene in unsaturated porous media*, Environmental Science and Technology, vol. 26(5), 1992.

Slichter, C.S., *Theoretical Investigation of the Motion of Ground Waters*, Originally published in U.S. Geological Survey Annual Report (1899), Ground Water, vol. 23, pp. 396-397, 1995.

Unger, A., Sudicky, E.A., and Forsyth, P., *Vacuum Extraction-Air Sparging DNAPL Remediation*, Water Resour. Res., vol. 31(8), pp. 1913-1925, 1995.

Rehfeldt, K.R., Boggs, J.M., and Gelhar, L.W., *Field Study of Dispersion in a Heterogeneous Aquifer, 3. Geostatistical Analysis of Hydraulic Conductivity*, Water Resources Research, vol. 28(12), pp. 3309-3324, 1992.

Sudicky, E.A., *A Natural Gradient Experiment on Solute Transport in a Sand and Gravel Aquifer: Spatial Variability and Its Role in the Dispersion Process*, Water Resources Research, vol. 22(13), pp. 2069-2082, 1986.

United States 1996 Budget, *Proposed Environmental Spending*, World Wide Web, http://www.doc.gov/cgi-bin/web_doc/public/bud96/Environment, 1995.

Voss, C., *A Finite-Element Simulation Model for Saturated-Unsaturated Fluid Density Dependent Ground Water Flow with Energy Transport or Chemically Reactive Single-Species Solute Transport*, USGS Water-Resources Investigations Report 84-4369, Washington, GPO, 1984.

Wolf, F., *Groundwater Site Characterization: A Systems Perspective*, Journal of Environmental Health, vol. 56(7), pp. 13-18, 1994.

Yeh, W.W., *Review of Parameter Identification Procedures in Groundwater Hydrology: The Inverse Problem*, Water Resources Research, vol. 22(2), pp. 95-108, 1986.

Zadeh, L.A., *The birth and evolution of fuzzy logic*, International Journal of General Systems, vol. 17, pp. 95-105, 1990.

VITA

Orlando José Doña, Jr. w[redacted]. A distinguished graduated of the University Laboratory School in 1978, he entered military service in the United States Army in the same year. After completing four years of honorable service in the regular army and two years in the army reserve, he was nominated as a United States Postal Service Letter Carrier in 1984. During this service, he begun his undergraduate education in Mathematics at San Diego State University in San Diego, California. He graduated summa cum laude in 1993 with a Bachelors degree in Mathematics, and received his commission on June 11, 1993 after having successfully completed the United States Air Force Reserve Officer Training Corps program.

Upon entering the active duty Air Force and after completing the Basic Communications Officer Training course at Keesler Air Force Base as a distinguished graduate, his first assignment was to Scott Air Force Base, Illinois. Assigned to the Communications Systems Program Office, Air Force Materiel Command, he served as Program Manager for the successful Base Information Digital Distribution System (BIDDS) Management Subsystem (BMS), In May of 1995, he entered the Graduate School of Engineering, Air Force Institute of Technology.

[redacted]